

The American Economic Review

914416

THE AMERICAN ECONOMIC ASSOCIATION

Founded in 1885

- Published at George Banta Co., Inc., Menasha, Wisconsin.

- THE AMERICAN ECONOMIC REVIEW, including four quarterly numbers, the *Proceedings* of the annual meetings, and *Directory* and Supplements, is published by the American Economic Association and is sent to all members five times a year, in March, May, June, September and December.

- Membership dues of the Association are \$20.00 a year, which includes a year's subscription to both the *American Economic Review* and the *Journal of Economic Literature*. Subscriptions by nonmembers are \$20.00 a year, and only subscriptions to both publications will be accepted. Single copies of the *Review* and *Journal* are \$4.00 each. Each order for copies of either publication must also include a \$.50 per order service charge. Orders should be sent to the Secretary's office, Nashville, Tennessee.

- Correspondence relating to the *Paper and Proceedings*, the *Directory*, advertising, permission to quote, business matters, subscriptions, membership and changes of address may be sent to the secretary, Rendigs Fels, 1313 21st Avenue, South, Nashville, Tennessee 37212. To be effective, notice of change of address must reach the secretary by the 1st of the month previous to the month of publication.

- Second-class postage paid at Nashville, Tennessee and at additional mailing offices. Printed in U.S.A.

Officers

President

JAMES TOBIN
Yale University

President-Elect

JOHN KENNETH GALBRAITH
Harvard University

Vice-Presidents

JAMES M. BUCHANAN
Virginia Polytechnic Institute
FRANCO MODIGLIANI
Massachusetts Institute of Technology

Secretary-Treasurer and Editor of Proceedings

RENDIGS FELS
Vanderbilt University

Managing Editor of The American Economic Review

GEORGE H. BORTS
Brown University

Managing Editor of The Journal of Economic Literature

MARK PERLMAN
University of Pittsburgh

Executive Committee

Elected Members of the Executive Committee

MARY JEAN BOWMAN
University of Chicago
CHARLES L. SCHULTZE
Brookings Institution
ROBERT DORFMAN
Harvard University
ARNOLD C. HARBERGER
University of Chicago
ROBERT EISNER
Northwestern University
JOHN R. MEYER
Yale University

Ex Officio Members

WILLIAM FELLNER
Yale University
WASSILY LEONTIEF
Harvard University

Cuk-402166-63-P014416

THE AMERICAN ECONOMIC REVIEW



March 1971

VOLUME LXI, NUMBER 1

GEORGE H. BORTS

Managing Editor

WILMA ST. JOHN

Assistant Editor

Board of Editors

BARBARA R. BERGMANN

JAGDISH N. BHAGWATI

PHILLIP CAGAN

GREGORY C. CHOW

CARL F. CHRIST

C. E. FERGUSON

HARRY G. JOHNSON

DANIEL MCFADDEN

ALVIN L. MARTY

HERBERT MOHRING

MARC NERLOVE

EDMUND S. PHELPS

G. WARREN NUTTER

VERNON L. SMITH

• Manuscripts and editorial correspondence relating to the regular quarterly issue of this REVIEW should be addressed to George H. Borts, Managing Editor of THE AMERICAN ECONOMIC REVIEW, Brown University, Providence, R.I. 02912. Manuscripts should be submitted in duplicate and in acceptable form. *Style Instructions* for guidance in preparing manuscripts will be provided upon request to the editor.

• No responsibility for the views expressed by authors in this REVIEW is assumed by the editors or the publishers, The American Economic Association.

• Copyright American Economic Association 1971.

Articles

- Theoretical Assumptions and Nonobserved Facts
Wassily Leontief 1

- Optimal Taxation and Public Production:
I—Production Efficiency
Peter A. Diamond and James A. Mirrlees 8

- United States Imports and Internal Pressure of
Demand: 1948–68
R. G. Gregory 28

- Cropsharing Tenancy in Agriculture: A Theoretical
and Empirical Analysis
P. K. Bardhan and T. N. Srinivasan 48

- On the Theory of the Competitive Firm Under
Price Uncertainty
Agnar Sandmo 65

- The Effect of Tariffs on Production, Consumption,
and Trade: A Revised Analysis
J. Clark Leith 74

- A General Disequilibrium Model of Income and
Employment
Robert J. Barro and Herschel I. Grossman 82

- A Test for Relative Efficiency and Application to
Indian Agriculture
Lawrence J. Lau and Pan A. Yotopoulos 94

- The Incidence of Social Security Payroll Taxes
John A. Brittain 110

- Determinants of the Commodity Structure of U.S.
Trade
Robert E. Baldwin 126

- Optimal Restrictions on Foreign Trade and Investment
Franz Gehrels 147

Communications

Large Industrial Corporations and Asset Shares:

Comment	<i>David R. Kamerschen</i>	160
Comment	<i>Stanley E. Boyle</i>	163
Reply	<i>David Mermelstein</i>	168

Welfare Aspects of a Regulatory Constraint: Note

<i>Eylan Sheshinski</i>	175
-------------------------	-----

Pitfalls in Financial Model Building: Some Extensions

<i>Mark L. Ladenson</i>	179
-------------------------	-----

Clothing Exemptions and Sales Tax Regressivity: Note

<i>David G. Davies</i>	187
------------------------	-----

An Economic Theory of the Second Moments of Disturbances of Behavioral Equations

<i>Henri Theil</i>	190
--------------------	-----

A Neglected Social Cost of a Voluntary Military

<i>Thomas E. Borcharding</i>	195
------------------------------	-----

On the Extension of Input-Output Analysis to Account for Environmental Externalities

<i>A. O. Converse</i>	197
-----------------------	-----

Mishan on the Gains from Trade:

Comment	<i>Mel Krauss and David M. Winch</i>	199
Reply	<i>E. J. Mishan</i>	202

Profit Constrained Revenue Maximization: Note

<i>Richard Rosenberg</i>	208
--------------------------	-----

Behavior of the Firm Under Regulatory Constraint: Note

<i>Israel Pressman and Arthur Carol</i>	210
---	-----

Spectral Analysis and the Detection of Lead-Lag Relations

<i>John C. Hause</i>	213
----------------------	-----

Subsidized Housing in a Competitive Market:

Comment	<i>Gordon Tullock</i>	218
Reply	<i>Edgar O. Olsen</i>	220

Expectations and the Demand for Bonds:

Comment	<i>Richard Roll</i>	225
Comment	<i>A. Buse</i>	229
Comment	<i>Reuben A. Kessel</i>	231
Reply	<i>John H. Wood</i>	232

Output of the Restrained Firm:

Comment	<i>A. Ross Shepherd</i>	237
Reply	<i>Milton Z. Kafoglis</i>	240

Production Indeterminacy with Three Goods and Two Factors:

A Comment on the Pattern of Trade	<i>Douglas B. Stewart</i>	241
Reply	<i>James R. Melvin</i>	245

In Memoriam: JACOB VINER

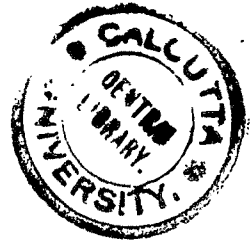
247

Errata

248

Notes

250



THE AMERICAN ECONOMIC REVIEW

VOL. LXI

MAY, 1971

NUMBER 2

PAPERS AND PROCEEDINGS

OF THE

Eighty-third Annual Meeting

OF THE

AMERICAN ECONOMIC ASSOCIATION

Detroit, Michigan, December 28-30, 1970

Edited by RENDIGS FELS, *Secretary of the Association*
with the assistance of

G. W. CHURCHILL, MARJORIE CHURCHILL, *and* MARTHA C. DRENNAN

Copyright American Economic Association, 1971



PRINTED BY GEORGE BANTA COMPANY, INC

Publication Office: Menasha, Wisconsin 54952

Executive Office: American Economic Association
1313 21st Avenue South, Nashville, Tennessee 37212

Inquiries and other communications regarding membership, meetings, and the general affairs of the Association, as well as orders for the publications and permission to reprint, should be addressed to Rendigs Fels, Secretary of the American Economic Association, 1313 21st Avenue South, Nashville, Tennessee 37212.

Changes of address should also be sent to the Secretary at the address given above. To be effective, notice must reach the office by the 1st of the month previous to the month of publication. The Association cannot be responsible for copies lost due to failure to report change of address in time.

Second-class postage paid at Nashville, Tennessee and at additional mailing offices.

No responsibility for the views expressed by authors in this *Review* is assumed by the editors or the publishers, the American Economic Association.

Membership dues of the Association are \$20.00 a year, \$16.00 for each year's dues being in payment of a year's subscription to both the *American Economic Review* and the *Journal of Economic Literature*. The *American Economic Review* is published five times a year with four quarterly issues, the *Papers and Proceedings* of the Annual Meetings, and supplements, and is sent to all members of the American Economic Association as one of the privileges of membership, as is the *Journal of Economic Literature* (four quarterly issues per year). Subscriptions by nonmembers are \$20.00 a year, and only subscriptions to both publications will be accepted. Single copies of the *Review* and *Journal* are \$4.00, May *Papers and Proceedings*, \$8.00, and 1969 *Handbook*, \$10.00 each. Each order for copies of the publications must include a \$.50 per order service charge. Orders should be sent to the Executive Office. Printed in U.S.A.

PRICE \$8.00

TABLE OF CONTENTS

Program of the Eighty-third Annual Meeting	vii
--	-----

PAPERS

Richard T. Ely Lecture The Keynesian Revolution and the Monetarist Counter-Revolution <i>Harry G. Johnson</i>	1	Conglomerate and Vertical Responses to Market Imperfection Expectations, Mergers and Equilib- rium in Purely Competitive Securi- ties Markets <i>John Lintner</i>	101
Current Status of Income Maintenance Experiments The Graduated Work Incentive Ex- periments: Current Progress <i>Har- old W. Watts</i>	15	The Vertical Integration of Produc- tion: Market Failure Considera- tions <i>Oliver E. Williamson</i>	112
The Seattle Experiment: The Com- bined Effect of Income Mainte- nance and Manpower Investments <i>Mordecai Kurz</i> and <i>Robert G. Spiegelman</i>	22	Discussion <i>Roland N. McKean</i>	124
The Gary Income Maintenance Ex- periment: Plans and Progress <i>Ter- ence F. Kelly</i> and <i>Leslie Singer</i> ..	30	<i>J. Fred Weston</i>	125
Discussion <i>James N. Morgan</i>	39	Micro Aspects of Macro Performance The Impact of Inflation on the Dis- tribution of Income and Wealth <i>Edward C. Budd</i> and <i>David F. Seiders</i>	128
The State of Economics: The Behavioral and Social Sciences Survey Introductory Remarks <i>Harry W. Riecken</i>	43	The Two-Way Relationship between the Budget and Economic Vari- ables <i>Arnold H. Packer</i>	139
The Reviewers Reviewed <i>Charles L. Schultze</i>	45	Discussion <i>Charles E. Metcalf</i>	150
The State of Political Economics <i>John G. Gurley</i>	53	<i>John L. Palmer</i>	150
Discussion <i>Robert M. Solow</i>	63	The Political Economy of Environmental Quality Environmental Pollution: Economics and Policy <i>Allen V. Kneese</i>	153
<i>Robert L. Heilbroner</i>	65	Discussion <i>Kenneth E. Boulding</i>	167
<i>Henry W. Riecken</i>	67	<i>Jack W. Carlson</i>	169
Television: Old Theories, Current Facts, and Future Policies The Growth of Cable TV and Its Probable Impact on Over-The-Air Broadcasting <i>Rolla Edward Park</i> ..	69	<i>Richard W. Judy</i>	172
Concentration of Control and the Price of Television Time <i>John L. Peterman</i>	74	<i>Marc J. Roberts</i>	173
Program Duplication, Diversity, and Effective Viewer Choices: Some Empirical Findings <i>Harvey J. Levin</i>	81	Commercial Policy and Less Developed Countries Trade Policies in Developing Coun- tries <i>Bela Balassa</i>	178
TV Program Diversity—New Evi- dence and Old Theories <i>Edward Greenberg</i> and <i>Harold J. Barnett</i> ..	89	Tariff Preferences and Separable Utility <i>Christopher Clague</i>	188
Discussion <i>John J. McGowan</i> and <i>Merton J. Peck</i>	94	Nontariff Barriers and the Export Performance of Developing Econ- omies <i>Ingo Walter</i>	195
		Discussion <i>Harald B. Malmgren</i>	206
		<i>Stephen Guisinger</i>	207
		<i>Robert E. Baldwin</i>	209
		Evaluation of Economic Regulation of Industry Methodology of Evaluating Economic Regulation <i>Fred M. Westfield</i> ..	211

Costs and Benefits of Regulating Communications <i>Boyd L. Nelson</i> ..	218	Market Choices and Optimum City Size <i>Edwin S. Mills and David M. de Ferranti</i>	340
The Social Costs of Regulating the Railroads <i>Ann F. Friedlaender</i> ..	226	Social Problems and the Urban Crisis: Can Public Policy Make a Difference? <i>Worth Bateman and Harold M. Hochman</i>	346
Discussion		Discussion	
<i>Alfred E. Kahn</i>	235	<i>Roland Artle</i>	354
<i>Larry F. Darby</i>	237	<i>Stephen P. Dresch</i>	355
<i>James R. Nelson</i>	239	<i>Anthony H. Pascal</i>	357
Teaching Economics: Experiments and Results		The Allocation of Social Risk	
The Lasting Effects of Elementary Economics Courses: Some Preliminary Results <i>Phillip Saunders</i> ..	242	Social Risk and Financial Markets <i>William Brainard and F. T. Dolbear</i>	360
Teaching Economics to Black Students <i>R. Grann Lloyd</i>	249	The Role of Insurance in the Allocation of Risk <i>Richard Kihlstrom and Mark Pauly</i>	371
Videotaped Dialogues in Economics <i>Barry Castro</i>	256	Insurance, Information, and Individual Action <i>Michael Spence and Richard Zeckhauser</i>	380
Discussion		Discussion	
<i>Irving Morrissett</i>	260	<i>Martin J. Bailey</i>	388
<i>Marcia L. Halvorsen</i>	262	<i>Hyman P. Minsky</i>	389
<i>Campbell R. McConnell</i>	264	<i>Donald D. Hester</i>	390
Educational Production Relationships		Population and Environment in the United States	
Education, Technology, and the Characteristics of Worker Productivity <i>Herbert Gintis</i>	266	Impact of Population on Resources and the Environment <i>Joseph L. Fisher</i>	392
Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data <i>Eric Hanushek</i> ..	280	Does Human Fertility Adjust to the Environment? <i>Richard A. Easterlin</i>	399
Ability and Schooling as Determinants of Lifetime Earnings or If You're So Smart, Why Aren't You Rich? <i>John C. Hause</i>	289	Issues in the Economics of a Population Policy for the United States <i>Glen G. Cain</i>	408
Discussion		Discussion	
<i>William T. Garner</i>	299	<i>T. Paul Schultz</i>	418
<i>Roger Alcaly</i>	301	<i>Paul Demeny</i>	419
<i>Charles Z. Wilson</i>	302		
Round Table Session on the Academic Labor Market		The Theory of Planning	
Whither the Market for Academic Economists? <i>Allan M. Cartter</i> ..	305	Decentralized Planning Procedures and Centrally Planned Economies <i>Richard D. Portes</i>	422
The Supply of Economists in the 1970's <i>Lindsey R. Harmon</i>	311	The Hungarian Economic Reform, Past and Future <i>Tamas Nagy</i> ..	430
Recent Behavior of Economists' Salaries <i>Francis M. Boddy</i>	316	Discussion	
Improving the Operation of the Academic Labor Market <i>David G. Brown</i>	321	<i>Edward Ames</i>	436
A Younger Economist's Views of the Market <i>Robert P. Strauss</i>	327	<i>Bela Balassa</i>	437
Urban Growth and Development		The Economics of Political Decentralization	
Jurisdictional Fragmentation and Residential Choice <i>Bryan Ellickson</i> ..	334	Towards a Predictive Theory of Inter-	

CONTENTS

v

governmental Grants <i>David F. Bradford</i> and <i>Wallace E. Oates</i> ..	440	Finance Committee <i>C. Wells Farnham</i>	487
State and Local Response to Fiscal Decentralization <i>A. James Heins</i>	449	Auditor <i>Arthur Andersen & Co</i>	490
Centralization and Decentralization: The Political Economy of Public School Systems <i>O. A. Davis</i> and <i>K. O. Kortanek</i>	456	Managing Editor, <i>American Economic Review</i> <i>George H. Borts</i>	498
Discussion		Managing Editor, <i>Journal of Economic Literature</i> <i>Mark Perlman</i>	503
<i>Charles J. Goetz</i>	463	Committee on Economic Education <i>G. L. Bach</i>	505
<i>Thomas E. Borcharding</i>	463	Visiting Scientists Program in Economics <i>Phillip Saunders</i>	507
<i>Walter E. Williams</i>	465	Representative to the National Bureau of Economic Research <i>Willard L. Thorp</i>	508
PROCEEDINGS		Representative to the International Economic Association <i>Fritz Machlup</i>	511
Minutes of the Business Meeting	471	Institute Policy and Advisory Board <i>Henry Rosovsky</i>	512
Minutes of the Executive Committee Meetings	474	Census Advisory Committee <i>Bert C. Hickman</i>	513
Reports		Economic Index Project <i>Richard Ruggles</i>	514
Secretary <i>Rendigs Fels</i>	481		
Treasurer <i>Rendigs Fels</i>	486		

PROGRAM OF THE EIGHTY-THIRD ANNUAL MEETING OF THE
AMERICAN ECONOMIC ASSOCIATION

Detroit, Michigan, December 27-30, 1970

Sunday, December 27, 1970

2:00 P.M.

Executive Committee Meeting

Monday, December 28, 1970

8:30 A.M.

Round Table Session on the Academic Labor Market

Chairman: W. LEE HANSEN, University of Wisconsin

Papers: LINDSEY R. HARMON, National Research Council; FRANCIS M. BODDY, University of Minnesota; DAVID G. BROWN, Miami University; ROBERT STRAUSS, University of North Carolina; ALLAN M. CARTTER, New York University

Television: Old Theories, Current Facts and Future Policies

Chairman: PETER O. STEINER, University of Michigan

Papers: ROLLA EDWARD PARK, Rand Corporation; JOHN PETERMAN, University of Chicago; HARVEY J. LEVIN, Hofstra University; EDWARD GREENBERG and HAROLD BARNETT, Washington University

Discussion Paper: MERTON J. PECK and JOHN J. MCGOWAN, Yale University

The Rising Cost of Health Care: Causes and Potential Remedies (Joint session with the Health Economics Research Organization)

Chairman: DONALD E. YETT, University of Southern California

Papers: DAVID S. SALKEVER, U.S. Public Health Service; VICTOR R. FUCHS, City University of New York; MARCIA J. KRAMER, National Bureau of Economic Research; HERBERT E. KLARMAN, New York University

Discussants: RALPH ANDREANO, University of Wisconsin; SYLVESTER E. BERKI, University of Michigan; MICHAEL INTRILIGATOR, University of Southern California; DONALD E. YETT, University of Southern California

Issues in Grants Economics (Symposium on Grants Economics 1) (Joint session with the Association for the Study of the Grants Economy)

Chairman: KENNETH E. BOULDING, University of Colorado

Papers: MARTIN PFAFF and ANITA B. PFAFF, Wayne State University; HENRY AARON, University of Maryland and the Brookings Institution; GEORGE VON FURSTENBERG, Indiana University; DAVID O. PORTER, University of California, Riverside; DAVID C. WARNER, Wayne State University; MURRAY L. WEIDENBAUM, U.S. Treasury Department and Washington University

10:30 A.M.

Urban Growth and Development

Chairman: JOHN F. KAIN, Harvard University

Papers: BRYAN ELLICKSON, University of California, Los Angeles; EDWIN S. MILLS, Princeton University; HAROLD M. HOCHMAN and C. WORTH BATEMAN, The Urban Institute

Discussants: ROLAND ARTLE, University of California, Berkeley; STEPHEN P. DRESCH, National Bureau of Economic Research; ANTHONY H. PASCAL, The Rand Corporation

The Allocation of Social Risk (Joint session with the Econometric Society)

Chairman: JACOB MARSCHAK, University of California, Los Angeles

Papers: WILLIAM BRAINARD, Yale University, and F. TRENERY DOLBEAR, Jr., Brandeis University; MARK PAULY and RICHARD KIELSTROM, Northwestern University; RICHARD ZECKHAUSER and MICHAEL SPENCE, Harvard University

Discussants: MARTIN BAILEY, University of Rochester; HYMAN MINSKY, Washington University; DONALD D. HESTER, University of Wisconsin

Statistical Appraisals of Recent Short and Long Term Monetary and Fiscal Policy (Joint session with the American Statistical Association)

Chairman: THOMAS MAYER, University of California, Davis

Papers: THOMAS F. CARGILL, Purdue University; VITTORIO BONOMO and CHARLES SCHOTTA, Virginia Polytechnic Institute; KEITH M. CARLSON, Federal Reserve Bank of St. Louis

Discussants: DAVID FAND, Wayne State University; WILLIAM YOHE, Duke University; MARK WILLES, Federal Reserve Bank of Philadelphia

2:30 P.M.

Population and Environment in the United States

Chairman: JOSEPH J. SPENGLER, Duke University

Papers: JOSEPH L. FISHER, Resources for the Future; RICHARD A. EASTERLIN, University of Pennsylvania; GLEN G. CAIN, University of Wisconsin

AMERICAN ECONOMIC ASSOCIATION

Discussants: T. PAUL SCHULTZ, The Rand Corporation; PAUL DEMENY, University of Hawaii; ROGER REVELLE, Harvard University
Evaluation of Economic Regulation of Industry (Joint session with the Transportation and Public Utilities Group of the AEA)
Chairman: GEORGE W. WILSON, Indiana University
Papers: FRED M. WESTFIELD, Vanderbilt University; BOYD NELSON, Federal Communications Commission; ANN FRIEDLAENDER, Boston College
Discussants: ALFRED E. KAHN, Cornell University; LARRY DARBY, Temple University; JAMES R. NELSON, Amherst College

2:30 P.M.

Financial Constraints and Minority Economic Development (Joint session with the American Finance Association)

Chairman: ANDREW F. BRIMMER, Board of Governors of the Federal Reserve System
Papers: FRED E. CASE, University of California, Los Angeles; WILFRED J. GARVIN, Small Business Administration; HENRY S. TERRELL, Board of Governors of the Federal Reserve System; EDWARD D. IRONS, Howard University and the National Bankers Association; ANDREW F. BRIMMER, Board of Governors of the Federal Reserve System
Discussants: MARCUS ALEXIS, Northwestern University and University of California, Berkeley; WILLIAM G. GRIGSBY, University of Pennsylvania; ROBERT LINDSAY, JR., New York University

8:30 P.M.

Richard T. Ely Lecture

Chairman: JAMES TOBIN, Yale University
Speaker: HARRY G. JOHNSON, University of Chicago and London School of Economics

Tuesday, December 29, 1970

8:30 A.M.

The Political Economy of Environmental Quality

Chairman: ROBERT H. HAVEMAN, University of Wisconsin
Paper: ALLEN V. KNEESE, Resources for the Future, Inc.
Discussants: KENNETH BOULDING, University of Colorado; JACK W. CARLSON, U.S. Office of Management and Budget; RICHARD JUDY, University of Toronto; MARC J. ROBERTS, Harvard University

The Economics of Political Decentralization (Joint session with the Public Choice Society)

Chairman: JAMES M. BUCHANAN, Virginia Polytechnic Institute
Papers: DAVID F. BRADFORD and WALLACE OATES, Princeton University; JAMES HEINS, University of Illinois; KENNETH O. KORTANEK and OTTO A. DAVIS, Carnegie-Mellon University
Discussants: CHARLES J. GOETZ, Virginia Polytechnic Institute; THOMAS BORCHERDING, University of Washington; WALTER E. WILLIAMS, University of California, Los Angeles; JOHN G. HEAD, Dalhousie University

Graduate Student Papers (Joint session with Omicron Delta Epsilon)

Chairman: PROFESSOR ERVIN ZINGLER, University of Houston
Papers: JOHN R. DOMINGUEZ, Massachusetts Institute of Technology; DENNIS BREEDEN, University of Tennessee; WOLFGANG W. FRANZ, Central Washington State College
Discussants: ZUBAIR IGBAL, Michigan State University; MICHAEL SARRIS, Wayne State University; SYLVESTER SCHIEBER, University of Notre Dame

10:30 A.M.

Current Status of Income Maintenance Experiments (Joint session with the Econometric Society)

Chairman: ROBERT LEVINE, The Rand Corporation
Papers: HAROLD W. WATTS, University of Wisconsin; MORDECAI KURZ, Stanford University, and ROBERT G. SPIEGELMAN, Stanford Research Institute; LESLIE P. SINGER, Indiana University; JAMES MORGAN, University of Michigan; GUY ORCUTT, Yale University

Commercial Policy and Less Developed Countries

Chairman: LAWRENCE KRAUSE, The Brookings Institution
Papers: BELA BALASSA, Johns Hopkins University and International Bank for Reconstruction and Development; CHRISTOPHER CLAGUE, University of Maryland; INGO WALTER, New York University
Discussants: HARALD B. MALMGREN, Overseas Development Council; ROBERT BALDWIN, University of Wisconsin; STEPHEN GUISSINGER, Southern Methodist University

10:30 A.M.

Progress in Socioeconomic Accounting (Joint session with the American Statistical Association)

Chairman: OTIS DUDLEY DUNCAN, University of Michigan

PROGRAM OF THE EIGHTY-THIRD ANNUAL MEETING

ix

Papers: MANCUR OLSON, University of Maryland; RICHARD RUGGLES and NANCY RUGGLES, Yale University; ROSANNE COLE, International Business Machines Corporation

Discussants: ROBERT EISNER, Northwestern University; MILTON MOSS, Office of Statistical Policy, U.S. Office of Management and Budget

12:30 P.M.

Economic Outlook Luncheon (Joint session with the American Finance Association and the American Statistical Association)

Chairman: MARTIN R. GAINSBROUGH

Speaker: PAUL A. VOLCKER, Under Secretary of the Treasury

2:30 P.M.

Position and Prospects of Blacks in the National Economy (Joint session with the American Agricultural Economics Association)

Chairman: JAMES G. MADDOX, North Carolina State University

Papers: HENRY PONDER, Alabama Agricultural and Mechanical College; CALVIN L. BEALE, United States Department of Agriculture; GEORGE S. TOLLEY, University of Chicago; PHYLLIS WALLACE, Metropolitan Applied Research Center

The State of Economics: The Behavioral and Social Sciences Survey

Chairman: WASSILY LEONTIEF, Harvard University

Papers: CHARLES SCHULTZE, The Brookings Institution; JOHN G. GURLEY, Stanford University; HENRY RIECKEN, Social Science Research Council; ROBERT M. SOLOW, Massachusetts Institute of Technology; ROBERT HEILBRONER, New School for Social Research

Intra and Intergeneration Transfers (Symposium on Grants Economics III) (Joint session with the Association for the Study of the Grants Economy)

Chairman: MARTIN PFAFF, Wayne State University

Papers: FREDERICK L. PRYOR, Swarthmore College; BENJAMIN BRIDGES, U.S. Department of Health, Education, and Welfare; E. LIEFMANN-KEIL, University of the Saarland, West Germany; MOLLIE ORSHANSKY, U.S. Department of Health, Education, and Welfare

Educational Production Relationships

Chairman: FINIS WELCH, National Bureau of Economic Research

Papers: HERBERT GINTIS, Harvard University; ERIC A. HANUSHEK, United States Air Force Academy; JOHN C. HAUSE, University of Minnesota and National Bureau of Economic Research

Discussants: ALAN THOMAS, University of Chicago; ROGER ALCALY, Columbia University; CHARLES WILSON, University of California, Los Angeles

8:00 P.M.

Presidential Address

Chairman: THEODORE W. SCHULTZ, University of Chicago

Speaker: WASSILY LEONTIEF, Harvard University

9:15 P.M.

Business Meeting

Wednesday, December 30, 1970

8:30 A.M.

The Theory of Planning (Joint session with the Association for Comparative Economics, the Association for the Study of Soviet-Type Economies, and the Econometric Society)

Chairman: EVSEY D. DOMAR, Massachusetts Institute of Technology

Papers: MICHAEL MANOVE, University of Michigan; RICHARD D. PORTES, Princeton University; TAMAS NAGY, Institute of Economics, Hungarian Academy of Sciences

Discussants: J. MICHAEL MONTIAS, Yale University; EDWARD AMES, State University of New York, Stony Brook; BELA BALASSA, The Johns Hopkins University

Micro Aspects of Macro Performance

Chairman: ARTHUR M. OKUN, The Brookings Institution

Papers: EDWARD BUDD and DAVID SEIDERS, Pennsylvania State University and United States Department of Commerce; ARNOLD PACKER, U.S. Office of Management and Budget

Discussants: CHARLES METCALF, University of Wisconsin; JOHN PALMER, Stanford University

Economic Thought

Chairman: CRAUFURD D. GOODWIN, Duke University

Papers: KENNETH BOULDING, University of Colorado; WILLIAM GRAMPP, University of Illinois, Chicago; VINCENT TARASCIO, University of North Carolina, Chapel Hill

AMERICAN ECONOMIC ASSOCIATION

10:30 A.M.

The Vietnamese War: Economic Prospects After Demobilization (Joint session with the Econometric Society)

Chairman: DANIEL SUITS, University of California, Santa Cruz; BERNARD UDIS, University of Colorado

Discussants: MURRAY WEIDENBAUM, U.S. Treasury and Washington University; SANDOR FUCHS, University of California, Berkeley; LAWRENCE LYNN, Stanford University

Grant Elements in Soviet-Type Economies (Joint session with the Association for the Study of the Grants Economy)

Chairman: ROBERT W. CAMPBELL, Indiana University

Papers: PAUL JONAS, U.S. Agency for International Development, New Delhi, India; JAMES CARTER, University of Oregon; JANOS HORVATH, Butler University; ALAN W. BROWN, Indiana University

Conglomerate and Vertical Responses to Market Imperfection

Chairman: RICHARD CYERT, Carnegie-Mellon University

Papers: JOHN LINTNER, Harvard University; OLIVER WILLIAMSON, University of Pennsylvania

Discussants: ROLAND MCKEAN, University of Virginia; J. FRED WESTON, University of California, Los Angeles

New Directions in National Bureau Research I

Papers: WILLIAM LANDES, FINIS WELCH, and JOHN F. KAIN

2:30 P.M.

Teaching Economics: Experiments and Results

Chairman: RENDIGS FELS, Vanderbilt University

Papers: PHILLIP SAUNDERS, Indiana University; R. GRANN LLOYD, Tennessee State University; BARRY CASTRO, Hostos Community College

Discussants: IRVING MORRISETT, University of Colorado; MARCIA HALVORSEN, Spelman College; CAMPBELL MCCONNELL, University of Nebraska

Earnings Differentials

Chairman: MELVIN W. REDER, Stanford University

Papers: VICTOR R. FUCHS, City University of New York; LOWELL E. GALLAWAY, RICHARD K. VEDDER and GENE L. CHAPIN, Ohio University; WILLIAM R. BAILEY and ALBERT E. SCHWENK, Bureau of Labor Statistics

Discussants: THOMAS W. GAVETT, Bureau of Labor Statistics; MAHMOOD A. ZAIDI, University of Minnesota; GERALD W. SCULLY, Southern Illinois University

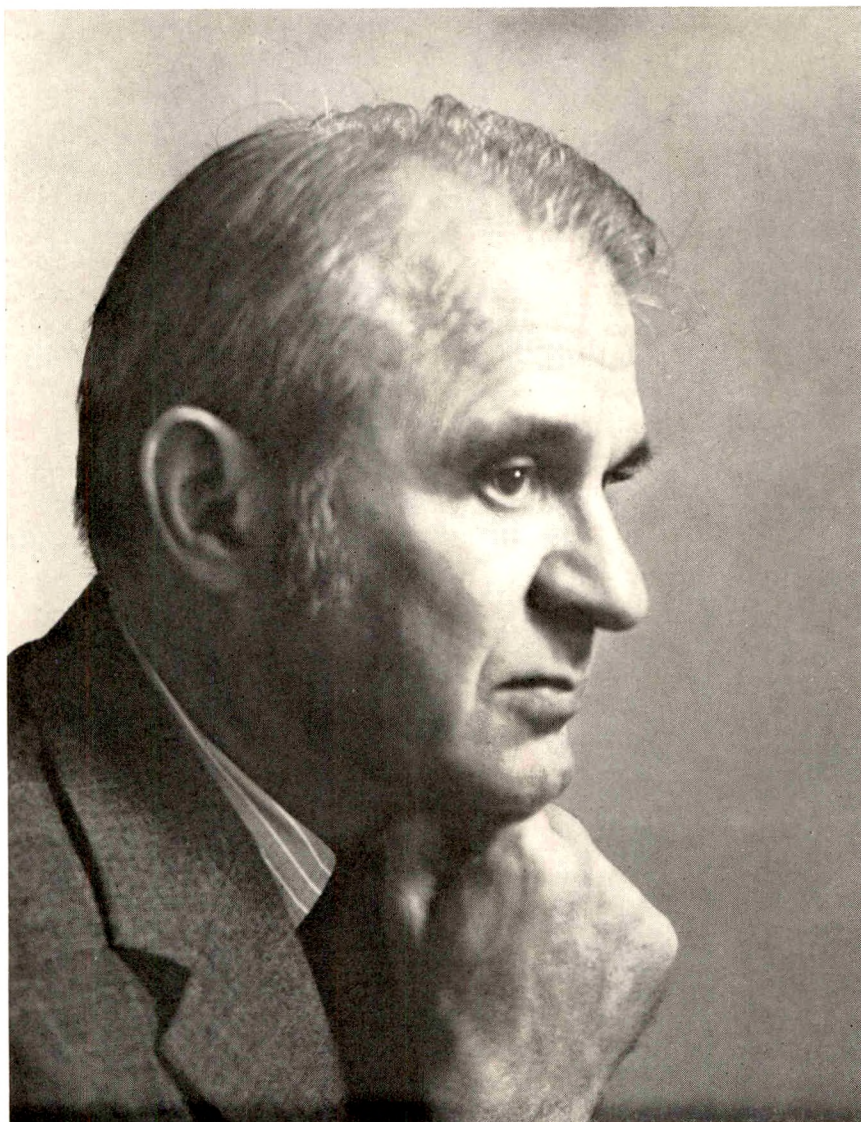
New Directions in National Bureau Research II (Joint session with the American Statistical Association and the National Bureau of Economic Research)

Papers: WARREN SANDERSON and ROBERT WILLIS, THOMAS JUSTER, and VICTOR ZARNOWITZ

THE purpose of the American Economic Association, according to its charter, is the encouragement of economic research, the issue of publications on economic subjects, and the encouragement of perfect freedom of economic discussion. The Association as such takes no partisan attitude, nor does it commit its members to any position on practical economic questions. It is the organ of no party, sect, or institution. Persons of all shades of economic opinion are found among its members, and widely different issues are given a hearing in its annual meetings and through its publications. The Association, therefore, assumes no responsibility for the opinions expressed by those who participate in its meetings. Moreover, the papers presented are the personal opinions of the authors and do not commit the organizations or institutions with which they are associated.

RENDIGS FELS
Secretary

Number 72 of a series of photographs of past presidents of the Association



Basil Beaton.

Theoretical Assumptions and Nonobserved Facts

By WASSILY LEONTIEF*

Economics today rides the crest of intellectual respectability and popular acclaim. The serious attention with which our pronouncements are received by the general public, hard-bitten politicians, and even skeptical businessmen is second only to that which was given to physicists and space experts a few years ago when the round trip to the moon seemed to be our only truly national goal. The flow of learned articles, monographs, and textbooks is swelling like a tidal wave; *Econometrica*, the leading journal in the field of mathematical economics, has just stepped up its publication schedule from four to six issues per annum.

And yet an uneasy feeling about the present state of our discipline has been growing in some of us who have watched its unprecedented development over the last three decades. This concern seems to be shared even by those who are themselves contributing successfully to the present boom. They play the game with professional skill but have serious doubts about its rules.

Much of current academic teaching and research has been criticized for its lack of relevance, that is, of immediate practical impact. In a nearly instant response to this criticism, research projects, seminars and undergraduate courses have been set up on poverty, on city and small town slums, on pure water and fresh air. In an almost Pavlovian reflex, whenever a new

complaint is raised, President Nixon appoints a commission and the university announces a new course. Far be it from me to argue that the fire should not be shifted when the target moves. The trouble is caused, however, not by an inadequate selection of targets, but rather by our inability to hit squarely any one of them. The uneasiness of which I spoke before is caused not by the *irrelevance* of the practical problems to which present day economists address their efforts, but rather by the palpable *inadequacy* of the scientific means with which they try to solve them.

If this simply were a sign of the overly high aspiration level of a fast developing discipline, such a discrepancy between ends and means should cause no worry. But I submit that the consistently indifferent performance in practical applications is in fact a symptom of a fundamental imbalance in the present state of our discipline. The weak and all too slowly growing empirical foundation clearly cannot support the proliferating superstructure of pure, or should I say, speculative economic theory.

Much is being made of the widespread, nearly mandatory use by modern economic theorists of mathematics. To the extent to which the economic phenomena possess observable quantitative dimensions, this is indisputably a major forward step. Unfortunately, any one capable of learning elementary, or preferably advanced calculus and algebra, and acquiring acquaintance with the specialized terminology of economics can set himself up as a theorist. Uncritical enthusiasm for mathematical formulation tends often to con-

* Presidential address delivered at the eighty-third meeting of The American Economic Association, Detroit, Michigan, December 29, 1970.

ceal the ephemeral substantive content of the argument behind the formidable front of algebraic signs.

Professional journals have opened wide their pages to papers written in mathematical language; colleges train aspiring young economists to use this language; graduate schools require its knowledge and reward its use. The mathematical model-building industry has grown into one of the most prestigious, possibly the most prestigious branch of economics. Construction of a typical theoretical model can be handled now as a routine assembly job. All principal components such as production functions, consumption and utility functions come in several standard types; so does the optional equipment as, for example, "factor augmentation"—to take care of technological change. This particular device is, incidentally, available in a simple exponential design or with a special automatic regulator known as the "Kennedy function." Any model can be modernized with the help of special attachments. One popular way to upgrade a simple one-sector model is to bring it out in a two-sector version or even in a still more impressive form of the "*n*-sector," that is, many-sector class.

In the presentation of a new model, attention nowadays is usually centered on a step-by-step derivation of its formal properties. But if the author—or at least the referee who recommended the manuscript for publication—is technically competent, such mathematical manipulations, however long and intricate, can even without further checking be accepted as correct. Nevertheless, they are usually spelled out at great length. By the time it comes to interpretation of the substantive *conclusions*, the assumptions on which the model has been based are easily forgotten. But it is precisely the empirical validity of these *assumptions* on which the usefulness of the entire exercise depends.

What is really needed, in most cases, is a very difficult and seldom very neat assessment and verification of these assumptions in terms of observed facts. Here mathematics cannot help and because of this, the interest and enthusiasm of the model builder suddenly begins to flag: "If you do not like my set of assumptions, give me another and I will gladly make you another model; have your pick."

Policy oriented models, in contrast to purely descriptive ones, are gaining favor, however nonoperational they may be. This, I submit, is in part because the choice of the final policy objectives—the selection and justification of the shape of the so-called objective function—is, and rightly so, considered based on normative judgment, not on factual analysis. Thus, the model builder can secure at least some convenient assumptions without running the risk of being asked to justify them on empirical grounds.

To sum up with the words of a recent president of the Econometric Society, "... the achievements of economic theory in the last two decades are both impressive and in many ways beautiful. But it cannot be denied that there is something scandalous in the spectacle of so many people refining the analysis of economic states which they give no reason to suppose will ever, or have ever, come about. . . . It is an unsatisfactory and slightly dishonest state of affairs."

But shouldn't this harsh judgment be suspended in the face of the impressive volume of econometric work? The answer is decidedly no. This work can be in general characterized as an attempt to compensate for the glaring weakness of the data base available to us by the widest possible use of more and more sophisticated statistical techniques. Alongside the mounting pile of elaborate theoretical models we see a fast-growing stock of equally intricate statistical tools. These

are intended to stretch to the limit the meager supply of facts.

Since, as I said before, the publishers' referees do a competent job, most model-testing kits described in professional journals are internally consistent. However, like the economic models they are supposed to implement, the validity of these statistical tools depends itself on the acceptance of certain convenient assumptions pertaining to stochastic properties of the phenomena which the particular models are intended to explain; assumptions that can be seldom verified.

In no other field of empirical inquiry has so massive and sophisticated a statistical machinery been used with such indifferent results. Nevertheless, theorists continue to turn out model after model and mathematical statisticians to devise complicated procedures one after another. Most of these are relegated to the stockpile without any practical application or after only a perfunctory demonstration exercise. Even those used for a while soon fall out of favor, not because the methods that supersede them perform better, but because they are new and different.

Continued preoccupation with imaginary, hypothetical, rather than with observable reality has gradually led to a distortion of the informal valuation scale used in our academic community to assess and to rank the scientific performance of its members. Empirical analysis, according to this scale, gets a lower rating than formal mathematical reasoning. Devising a new statistical procedure, however tenuous, that makes it possible to squeeze out one more unknown parameter from a given set of data, is judged a greater scientific achievement than the successful search for additional information that would permit us to measure the magnitude of the same parameter in a less ingenious, but more reliable way. This despite the fact that in all too many instances sophisti-

cated statistical analysis is performed on a set of data whose exact meaning and validity are unknown to the author or rather so well known to him that at the very end he warns the reader not to take the material conclusions of the entire "exercise" seriously.

A natural Darwinian feedback operating through selection of academic personnel contributes greatly to the perpetuation of this state of affairs. The scoring system that governs the distribution of rewards must naturally affect the make-up of the competing teams. Thus, it is not surprising that the younger economists, particularly those engaged in teaching and in academic research, seem by now quite content with a situation in which they can demonstrate their prowess (and incidentally, advance their careers) by building more and more complicated mathematical models and devising more and more sophisticated methods of statistical inference without ever engaging in empirical research. Complaints about the lack of indispensable primary data are heard from time to time, but they don't sound very urgent. The feeling of dissatisfaction with the present state of our discipline which prompts me to speak out so bluntly seems, alas, to be shared by relatively few. Yet even those few who do share it feel they can do little to improve the situation. How could they?

In contrast to most physical sciences, we study a system that is not only exceedingly complex but is also in a state of constant flux. I have in mind not the obvious change in the variables, such as outputs, prices or levels of employment, that our equations are supposed to explain, but the basic structural relationships described by the form and the parameters of these equations. In order to know what the shape of these structural relationships actually are at any given time, we have to keep them under continuous surveillance.

By sinking the foundations of our ana-

lytical system deeper and deeper, by reducing, for example, cost functions to production functions and the production functions to some still more basic relationships eventually capable of explaining the technological change itself, we should be able to reduce this drift. It would, nevertheless, be quite unrealistic to expect to reach, in this way, the bedrock of invariant structural relationships (measurable parameters) which, once having been observed and described, could be used year after year, decade after decade, without revisions based on repeated observation.

On the relatively shallow level where the empirically implemented economic analysis now operates even the more invariant of the structural relationships, in terms of which the system is described, change rapidly. Without a constant inflow of new data the existing stock of factual information becomes obsolete very soon. What a contrast with physics, biology or even psychology where the magnitude of most parameters is practically constant and where critical experiments and measurements don't have to be repeated every year!

Just to keep up our very modest current capabilities we have to maintain a steady flow of new data. A progressive expansion of these capabilities would be out of the question without a continuous and rapid rise of this flow. Moreover, the new, additional data in many instances will have to be qualitatively different from those provided hitherto.

To deepen the foundation of our analytical system it will be necessary to reach unhesitatingly beyond the limits of the domain of economic phenomena as it has been staked out up to now. The pursuit of a more fundamental understanding of the process of production inevitably leads into the area of engineering sciences. To

penetrate below the skin-thin surface of conventional consumption functions, it will be necessary to develop a systematic study of the structural characteristics and of the functioning of households, an area in which description and analysis of social, anthropological and demographic factors must obviously occupy the center of the stage.

Establishment of systematic cooperative relationships across the traditional frontiers now separating economics from these adjoining fields is hampered by the sense of self-sufficiency resulting from what I have already characterized as undue reliance on indirect statistical inference as the principal method of empirical research. As theorists, we construct systems in which prices, outputs, rates of saving and investment, etc., are explained in terms of production functions, consumption functions and other structural relationships whose parameters are assumed, at least for arguments' sake, to be known. As econometricians, engaged in what passes for empirical research, we do not try, however, to ascertain the actual shapes of these functions and to measure the magnitudes of these parameters by turning up new factual information. We make an about face and rely on indirect statistical inference to derive the unknown structural relationships from the observed magnitudes of prices, outputs and other variables that, in our role as theoreticians, we treated as unknowns.

Formally, nothing is, of course, wrong with such an apparently circular procedure. Moreover, the model builder in erecting his hypothetical structures is free to take into account all possible kinds of factual knowledge and the econometrician in principle, at least, can introduce in the estimating procedure any amount of what is usually referred to as "exogenous" information before he feeds his pro-

grammed tape into the computer. Such options are exercised rarely and when they are, usually in a casual way.

The same well-known sets of figures are used again and again in all possible combinations to pit different theoretical models against each other in formal statistical combat. For obvious reasons a decision is reached in most cases not by a knock-out, but by a few points. The orderly and systematic nature of the entire procedure generates a feeling of comfortable self-sufficiency.

This complacent feeling, as I said before, discourages venturesome attempts to widen and to deepen the empirical foundations of economic analysis, particularly those attempts that would involve crossing the conventional lines separating ours from the adjoining fields.

True advance can be achieved only through an iterative process in which improved theoretical formulation raises new empirical questions and the answers to these questions, in their turn, lead to new theoretical insights. The "givens" of today become the "unknowns" that will have to be explained tomorrow. This, incidentally, makes untenable the admittedly convenient methodological position according to which a theorist does not need to verify directly the factual assumptions on which he chooses to base his deductive arguments, provided his empirical conclusions seem to be correct. The prevalence of such a point of view is, to a large extent, responsible for the state of splendid isolation in which our discipline nowadays finds itself.

An exceptional example of a healthy balance between theoretical and empirical analysis and of the readiness of professional economists to cooperate with experts in the neighboring disciplines is offered by Agricultural Economics as it developed in this country over the last fifty years. A

unique combination of social and political forces has secured for this area unusually strong organizational and generous financial support. Official agricultural statistics are more complete, reliable, and systematic than those pertaining to any other major sector of our economy. Close collaboration with agronomists provides agricultural economists with direct access to information of a technological kind. When they speak of crop rotation, fertilizers, or alternative harvesting techniques, they usually know, sometimes from personal experience, what they are talking about. Preoccupation with the standard of living of the rural population has led agricultural economists into collaboration with home economists and sociologists, that is, with social scientists of the "softer" kind. While centering their interest on only one part of the economic system, agricultural economists demonstrated the effectiveness of a systematic combination of theoretical approach with detailed factual analysis. They also were the first among economists to make use of the advanced methods of mathematical statistics. However, in their hands, statistical inference became a complement to, not a substitute for, empirical research.

The shift from casual empiricism that dominates much of today's econometric work to systematic large-scale factual analysis will not be easy. To start with, it will require a sharp increase in the annual appropriation for Federal Statistical Agencies. The quality of government statistics has, of course, been steadily improving. The coverage, however, does not keep up with the growing complexity of our social and economic system and our capability of handling larger and larger data flows.

The spectacular advances in computer technology increased the economists' potential ability to make effective analytical use of large sets of detailed data. The time

is past when the best that could be done with large sets of variables was to reduce their number by averaging them out or what is essentially the same, combining them into broad aggregates; now we can manipulate complicated analytical systems without suppressing the identity of their individual elements. There is a certain irony in the fact that, next to the fast-growing service industries, the areas whose coverage by the Census is particularly deficient are the operations of government agencies, both federal and local.

To place all or even the major responsibility for the collection of economic data in the hands of one central organization would be a mistake. The prevailing decentralized approach that permits and encourages a great number of government agencies, non-profit institutions and private businesses engaged in data gathering activities acquitted itself very well. Better information means more detailed information and detailed specialized information can be best collected by those immediately concerned with a particular field. What is, however, urgently needed is the establishment, maintenance and enforcement of coordinated uniform classification systems by all agencies, private as well as public, involved in this work. Incompatible data are useless data. How far from a tolerable, not to say, ideal state our present economic statistics are in this respect, can be judged by the fact that because of differences in classification, domestic output data cannot be compared, for many goods, with the corresponding export and import figures. Neither can the official employment statistics be related without laborious adjustments to output data, industry by industry. An unreasonably high proportion of material and intellectual resources devoted to statistical work is now spent not on the collection of primary information but on a frustrating and wasteful struggle

with incongruous definitions and irreconcilable classifications.

Without invoking a misplaced methodological analogy, the task of securing a massive flow of primary economic data can be compared to that of providing the high energy physicists with a gigantic accelerator. The scientists have their machines while the economists are still waiting for their data. In our case not only must the society be willing to provide year after year the millions of dollars required for maintenance of a vast statistical machine, but a large number of citizens must be prepared to play, at least, a passive and occasionally even an active part in actual fact-finding operations. It is as if the electrons and protons had to be persuaded to cooperate with the physicist.

The average American does not seem to object to being interviewed, polled, and surveyed. Curiosity, the desire to find out how the economic system (in which most of us are small gears, and some, big wheels) works might in many instances provide sufficient inducement for cooperation of this kind.

One runs up, of course, occasionally against the attitude that "what you don't know can't hurt you" and that knowledge might be dangerous: it may generate a desire to tinker with the system. The experience of these years seems, however, to have convinced not only most economists—with a few notable exceptions—but also the public at large that a lack of economic knowledge can hurt badly. Our free enterprise system has rightly been compared to a gigantic computing machine capable of solving its own problems automatically. But any one who has had some practical experience with large computers knows that they do break down and can't operate unattended. To keep the automatic, or rather the semi-automatic, engine of our economy in good working order we must not only understand the general

principles on which it operates, but also be acquainted with the details of its actual design.

A new element has entered the picture in recent years—the adoption of methods of modern economic analysis by private business. Corporate support of economic research goes as far back as the early 1920's when Wesley Mitchell founded the National Bureau. However, it is not this concern for broad issues of public policies or even the general interest in economic growth and business fluctuations that I have in mind, but rather the fast-spreading use of advanced methods of Operations Research and of so-called Systems' Analysis. Some of the standard concepts and analytical devices of economic theory first found their way into the curricula of our business schools and soon after that, sophisticated management began to put them into practice. While academic theorists are content with the formulation of general principles, corporate operations researchers and practical systems' analysts have to answer questions pertaining to specific real situations. Demand for economic data to be used in practical business planning is growing at an accelerated pace. It is a high quality demand: business users in most instances possess first-hand technical knowledge of the area to which the data they ask for refer. Moreover, this demand is usually "effective." Profit-making business is willing and able to pay the costs of gathering the information it wants to have. This raises the thorny question of public access to privately collected data and of the proper division of labor and coopera-

tion between government and business in that fast-expanding field. Under the inexorable pressure of rising practical demand, these problems will be solved in one way or another. Our economy will be surveyed and mapped in all its many dimensions on a larger and larger scale.

Economists should be prepared to take a leading role in shaping this major social enterprise not as someone else's spokesmen and advisers, but on their own behalf. They have failed to do this up to now. The Conference of Federal Statistics Users organized several years ago had business, labor, and many other groups represented among its members, but not economists as such. How can we expect our needs to be satisfied if our voices are not heard?

We, I mean the academic economists, are ready to expound, to any one ready to lend an ear, our views on problems of public policy: give advice on the best ways to maintain full employment, to fight inflation, to foster economic growth. We should be equally prepared to share with the wider public the hopes and disappointments which accompany the advance of our own often desperately difficult, but always exciting intellectual enterprise. This public has amply demonstrated its readiness to back the pursuit of knowledge. It will lend its generous support to our venture too, if we take the trouble to explain what it is all about.

REFERENCE

- F. H. Hahn, "Some Adjustment Problems," *Econometrica*, Jan. 1970, 38, 1-2.

Optimal Taxation and Public Production I: Production Efficiency

By PETER A. DIAMOND AND JAMES A. MIRRELES*

Theories of optimal production in a planned economy have usually assumed that the tax system can allow the government to achieve any desired redistribution of property.¹ On the other hand, some recent discussions of public investment criteria have tended to ignore taxation as a complementary method of controlling the economy.² Although lump sum transfers of the kind required for full optimality³ are not feasible today, commodity and income taxes can certainly be used to increase welfare.⁴ We shall therefore examine the maximization of social welfare using

both taxes and public production as control variables. In doing so, we intend to bring together the theories of taxation, public investment, and welfare economics.

There are two main results of the study: the demonstration of the desirability of aggregate production efficiency in a wide variety of circumstances provided that taxes are set at the optimal level; and an examination of that optimal tax structure. It is widely known that aggregate production efficiency is desired as one part of achieving a Pareto optimum. It is also widely known that when the desired Pareto optimum cannot be achieved, aggregate production efficiency may not be desirable. Our conclusion differs from these results in that production efficiency is desirable although a full Pareto optimum is not achieved. In the optimum position, the presence of commodity taxes implies that marginal rates of substitution are not equal to marginal rates of transformation. Furthermore, the absence of lump sum taxes implies that the income distribution is not the best that can be conceived. Yet, the presence of optimal commodity taxes will be shown to imply the desirability of aggregate production efficiency.

This result is similar to that derived by Marcel Boiteux, although he considered an economy where lump sum redistributions of income were possible. Boiteux also examined the optimal tax structure that was necessary for this result. The optimal tax structure for the case of a single consumer (or equivalently with lump sum redistribution) has also been examined by Frank

* The authors are at Massachusetts Institute of Technology and Nuffield College, Oxford, respectively. During some of the work, Diamond was at Churchill College, Cambridge and Nuffield College, Oxford and Mirrlees was at M.I.T. Earlier versions of this paper were given at Econometric Society winter meetings at Washington and Blaricum, 1967, at the University Social Science Council Conference, Kampala, Uganda, December 1968, and to the Game Theory and Mathematical Economics Seminar, Hebrew University, Jerusalem. The authors wish to thank M.A.H. Dempster, D. K. Foley, P. A. Samuelson, K. Shell, and participants in these seminars for helpful discussions on this subject, and referees for valuable comments. Diamond was supported in part by the National Science Foundation under grant GS 1585. The authors bear sole responsibility for opinions and errors.

¹ For a discussion of this literature, see Abram Bergson.

² For a survey of this literature, see Alan Prest and Ralph Turvey.

³ We wish to distinguish here between lump sum taxes, which may vary from individual to individual while being unaffected by the individual's behavior, and poll taxes which are the same for all individuals, or perhaps for all individuals within several large groups, distinguished perhaps by age, sex, or region.

⁴ For another study of the general equilibrium impact of taxation, which does not explore the optimality question, see Gerard Debreu (1954).

Ramsey and Paul Samuelson.⁵ Our results move beyond theirs in considering the problem of income redistribution together with that of raising revenue. Even in the absence of government revenue requirements, if lump sum redistribution is impossible, the government will want to use its excise tax powers to improve income distribution. It will subsidize and tax different goods so as to alter individual real incomes. The optimal redistribution by this method occurs when there is a balance between the equity improvements and the efficiency losses from further taxation.

The general situation we want to discuss is an economy in which there are many consumers, public and private production, public consumption, and many different kinds of feasible tax instruments. We think that it is easier to understand the problem if we present the analysis first for a single consumer, no public consumption, and only commodity taxation, although this case has little intrinsic interest. The main point of the paper is that the analysis of this special case carries over in the main to the general case.

The first two sections are devoted to this special case. In the first, the situation is portrayed geometrically (for a two-commodity world with no private production); in the second, production efficiency and conditions for the optimal taxes are derived by application of the calculus. The use of the calculus here and elsewhere is not perfectly rigorous for the usual reasons. These issues are taken up in Section IV. In the third section, we extend the analysis of production to an economy with many consumers, elucidating precise conditions under which production efficiency is desirable (and presenting certain exceptions).

Section IV provides a rigorous statement of the theorems. In the fifth section, we discuss briefly certain applications and extensions of the basic efficiency result.

A following paper, referred to here as Diamond-Mirrlees II, will appear in the June 1971 *Review*. In it we will examine the optimality rules for commodity taxes, for other taxes including income taxes, and for public consumption. We will also give a rigorous statement of conditions under which the first-order conditions obtained (heuristically) below are indeed necessary conditions.

I. One-Consumer Economy— Geometric Analysis

We begin by considering an economy with a single, price-taking consumer and two commodities. We assume, for the moment, that all production possibilities are controlled by the government. While there is no scope for redistribution of income in this economy, the government might need to raise revenue to cover losses if there are increasing returns to scale or if there are fixed expenditures (such as defense) and constant returns to scale. Alternatively, the technology might exhibit decreasing returns to scale, facing the government with the problem of disposing of a surplus if all transactions are carried out at market prices. The optimal solution to either raising or disposing of revenue is well known. A poll tax or subsidy, as the case may be, will permit the hiring of the needed resources and permit the economy to achieve a Pareto optimum, which, in a one-consumer economy, is equivalent to the maximization of the consumer's utility. While this is a reasonable possibility in a one-consumer economy, lump sum taxes varying from individual to individual do not seem feasible in a much larger economy. An identical problem of distributing a surplus among many people arises if it is desired to improve income distribution.

⁵ For a detailed history of analysis of this problem, see William Baumol and David Bradford. A summary and discussion of the work of Boiteux has been given by Jacques Drèze.

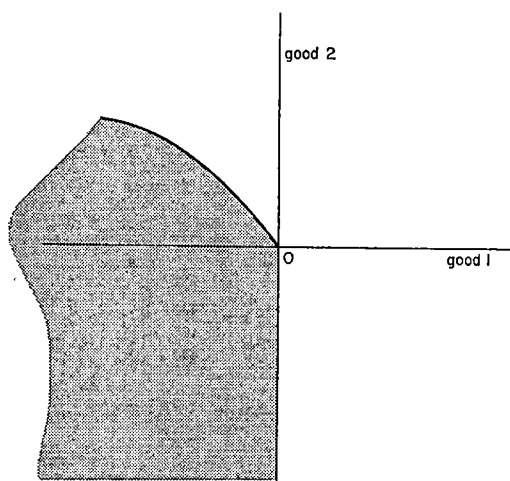


FIGURE 1

Thus we shall consider the use of commodity taxes when lump sum taxes are not permitted to the government, not for the intrinsic interest of this question in a one-consumer economy, but as an introduction to the many-consumer case. Furthermore we shall hold constant the government expenditure pattern which directly affects consumer utility. Thus we can ignore it, since the utility function already reflects its impact. The addition of choice for public consumption will be considered in Diamond-Mirrlees II.

Assuming free disposal, the technological constraint on the planner is that the government supply be on or under the production frontier. Such a constraint is shown by the shaded area in Figure 1. Let us measure on the axes the quantities supplied to the consumer. Thus, the output being produced (good 2) is measured positively, while the input (good 1) is measured negatively. The case drawn is the familiar one of decreasing returns to scale. If the government needed a fixed bundle of resources, for national defense say, then the production possibility frontier (describing the potential transactions with the consumer) would not pass through

the origin. With constant returns to scale this might appear as in Figure 2, where a units of good 1 are needed for defense. (It is perhaps convenient to think of good 1 as labor and good 2 as a consumption good.)

In a totally planned economy, where the planner selects a fixed consumption bundle (including labor to be supplied) for each consumer, the planner would have no further constraint and could choose any point that was technologically feasible. Again, this is not implausible for the planner in a one-consumer economy, but becomes so as the number of households grows. A more realistic assumption, then, is to assume that the planner can only deal with consumers through the market place, hiring labor and selling the consumer good. Assume further that the planner is constrained to charge uniform prices. The planner must now set the price of the consumer good relative to the wage (or inversely the real wage), and is constrained to transactions which the consumer is willing to undertake at some relative price. The locus of consumption bundles which the consumer is willing to achieve by trade from the origin is the offer curve or price-consumption locus. It represents the

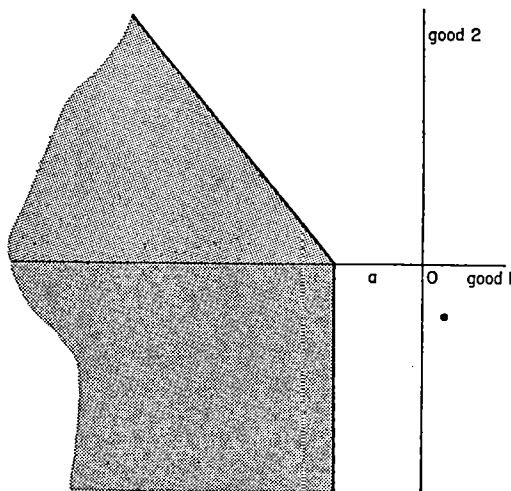


FIGURE 2

bundles of goods that the consumer would purchase at different possible price ratios. Figure 3 contains an example of an offer curve with several hypothetical budget lines and the corresponding indifference curves drawn in. The planner thus has two constraints: he must choose a point which is both technologically feasible and an equilibrium bundle from the point of view of the consumer. Combining these two constraints, the range of consumption bundles which are both feasible and potential consumer equilibria is shown as the heavy line in Figure 4.

We can state these two constraints algebraically. Let us denote by $z = (z_1, \dots, z_n)$ the vector of government supply. The production constraint is then written

$$(1) \quad G(z) \leq 0, \text{ or, equivalently,} \\ z_1 \leq g(z_2, z_3, \dots, z_n)$$

The constraint that the government sup-

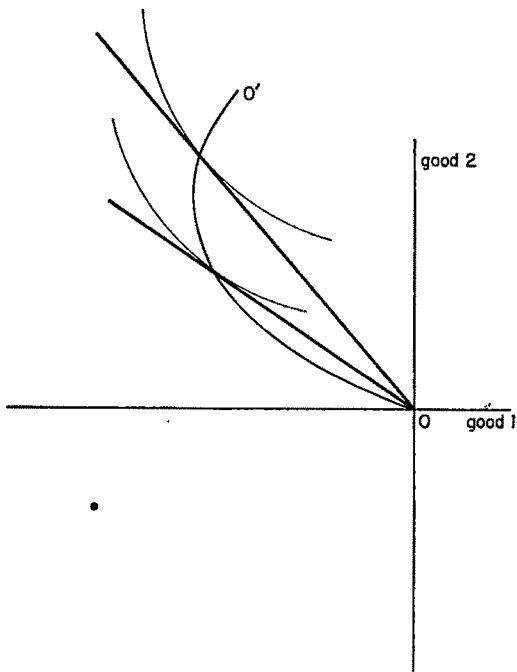


FIGURE 3

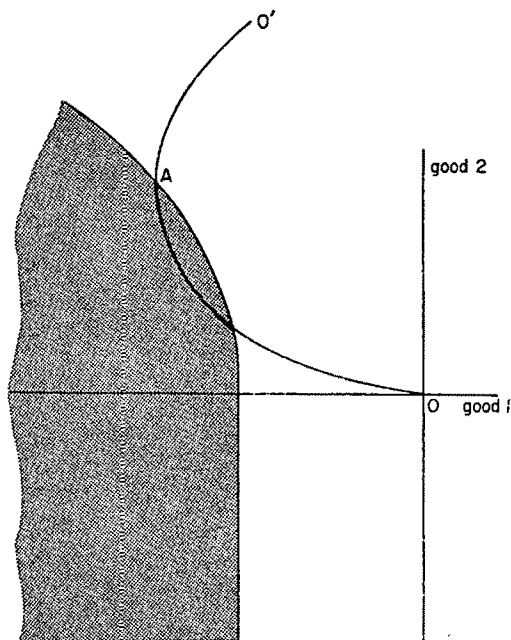


FIGURE 4

ply equal the consumer demand for some price can be written in vector notation

$$(2) \quad x(q) = z,$$

where $x = (x_1, \dots, x_n)$ is the vector of consumer demands and $q = (q_1, \dots, q_n)$ is the vector of prices faced by the consumer.

Now consider the government's objectives. Since the consumer's equilibrium position is determined by the prices he faces, we can, in the usual circumstances, describe the objective function as a function of prices, say $v(q)$. The problem is to choose q so as to

$$(3) \quad \begin{aligned} &\text{Maximize } v(q) \\ &\text{subject to } G(x(q)) \leq 0 \end{aligned}$$

This simply formulated problem is the focus of attention of the paper and can take on a variety of interpretations. The reader may note that the consideration of many consumers does not alter the form of this problem. This is a major advantage

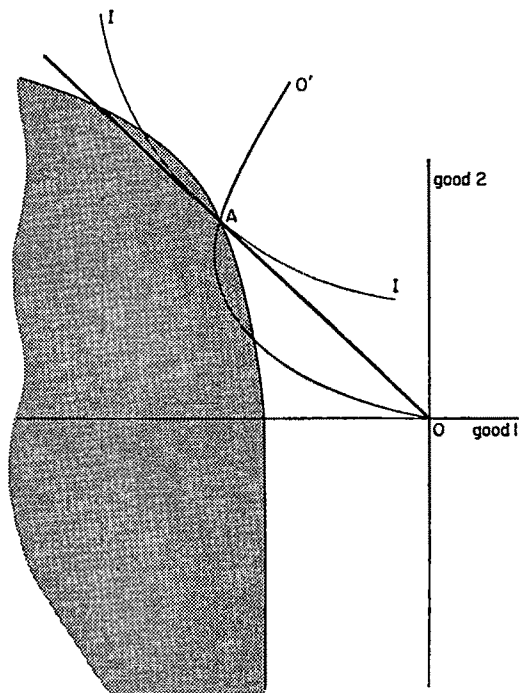


FIGURE 5

of using prices rather than quantities as the focus of the analysis.

Let us consider the case where the planner seeks to maximize the same function of consumption as the consumer's utility function. The welfare function is said to be *individualistic*, or to respect individual preferences, since welfare can be written as a function of individual utility. Returning to Figure 3 we see that the consumer moves to higher indifference curves as he proceeds along the offer curve away from the origin. Thus, in Figure 4 we wish to move as far along OO' as possible, subject to the constraint of the shaded production possibility set. The optimal point is therefore A , where the offer curve and the production frontier intersect.

The prices which will induce the consumer to purchase the optimal consumption bundle are defined by the budget line OA . In Figure 5 we show the optimal point and the implied budget line, and indiffer-

ence curve II . All the points above II and in the shaded production set are Pareto-superior to A and technologically feasible, but not attainable by market transactions without lump sum transfers. For contrast, in Figure 6, we show the Pareto optimal point, B , and the implied budget line, and indifference curve $I'I'$, which will permit decentralization. In the case drawn, the consumer's budget line does not pass through the origin; this represents his payment of a lump sum tax to cover government expenditures in excess of profits from production.

We see that the optimal point is on the production possibility frontier of the economy, not inside it. This important property of the optimum can easily be seen to carry over to the case of many commodities, but still one consumer. With many commodities, the offer curve is a union of loci, each of which is obtained by holding the prices of all but one commodity constant and varying the price of that one commodity. Doing this for each com-

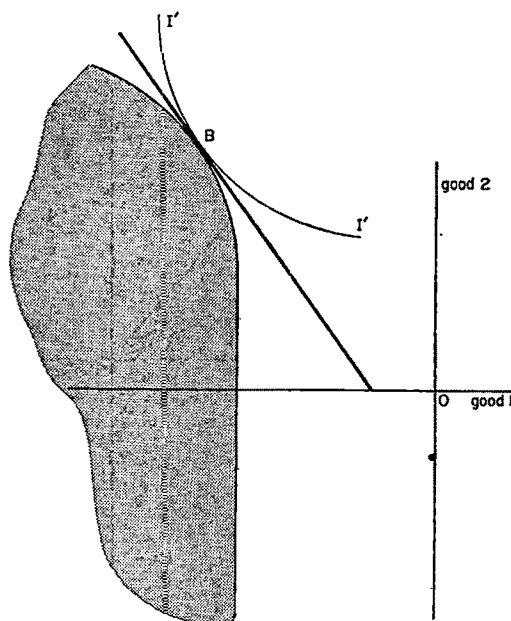


FIGURE 6

modity, and for all possible configurations of prices for the other commodities, generates all the loci. The offer curve is the union of such loci. On each locus, the point which is also on the production frontier is better than the other points on the locus. Thus, any point which is not on the production frontier is dominated by some point which is on the frontier. Therefore, the optimal point is one of the points on the frontier. The implications of this result will be seen more clearly below, when we consider both public and private production. For this result to carry over to the case of many consumers requires one further, mild assumption which will be discussed in the third section. First, we treat the one consumer economy algebraically, with both public and private production, showing by calculus the desirability of aggregate production efficiency, and obtaining the optimal relationship between consumer prices and the slope of the production possibilities. This relationship defines the optimal tax structure.

II. One-Consumer Economy— Algebraic Analysis

We assume constant returns to scale in the private production sector and the presence of competitive conditions there. In equilibrium there are, therefore, no profits. (This is a critical assumption for the efficiency analysis.) We also assume, for the present, that the only taxes used by the government are commodity taxes.⁶ Consumer prices, q , therefore determine the choices available to the consumer, and we may write the welfare function as a function of consumer prices, $v(q)$. Notice that this covers the case where the government's assessment of welfare does not coincide with the consumer's utility, al-

though depending on what he consumes. In the special case where social preferences coincide with those of the single consumer, his utility may be taken to measure welfare. Then we have

$$(4) \quad v(q) = u(x(q))$$

We shall not use this special form for $v(q)$ in the analysis below until we come to evaluate the tax structure explicitly. Until that point, the analysis applies also to welfare functions that are not individualistic. For later use let us express the derivatives of v in this special case. Writing $v_k = \partial v / \partial q_k$, $u_i = \partial u / \partial x_i$, and using (4), we have

$$(5) \quad v_k = \sum u_i \frac{\partial x_i}{\partial q_k} = -\alpha x_k,$$

where α is a positive constant (i.e., independent of k), the marginal utility of income. Equation (5) follows from the budget constraint,

$$(6) \quad \sum q_i x_i = 0,$$

which on differentiation with respect to q_k yields

$$(7) \quad x_k + \sum q_i \frac{\partial x_i}{\partial q_k} = 0$$

Since utility-maximization implies that $u_i = \alpha q_i$, (5) now follows from (7).

Production

Let us denote the vector of prices faced by private producers by $p = (p_1, \dots, p_n)$. Because of taxes, t , these may differ from the prices faced by consumers: $q_i = p_i + t_i$ ($i = 1, \dots, n$). $y = (y_1, \dots, y_n)$ is the vector of commodities privately supplied (inputs will thus appear as negative supplies), and we write the private production constraint,

$$(8) \quad y_1 = f(y_2, \dots, y_n)$$

Notice that we assume *equality* in the

⁶ This assumption is made solely for simplicity. In Diamond-Mirrlees II, the general principles will be seen to carry over with additional taxes, including a progressive income tax.

production constraint, that is, that production is efficient in the private sector. This follows from profit maximization if there are no zero prices. We assume that f is a differentiable function, and that $y_i \neq 0$ ($i=1, \dots, n$). Then, profit maximization means that

$$(9) \quad p_i = -p_1 f_i(y_2, \dots, y_n), \quad (i=2, \dots, n)$$

where f_i denotes the derivative of f with respect to y_i . Also, by the assumption of constant returns to scale, maximized profits are zero in equilibrium:

$$(10) \quad \sum p_i y_i = 0$$

So that we may conveniently employ calculus, we shall assume that the government production constraint, (1), is satisfied with an equality rather than an inequality:

$$(11) \quad z_1 = g(z_2, \dots, z_n)$$

Thus we do not give the government the option of inefficient government production. Rather, we shift our attention to *aggregate* production efficiency. Efficiency will be present if marginal rates of transformation are the same in publicly and privately controlled production. It will then be seen quite easily that the assumption of efficiency in the public sector is justified.

Walras' Law

We have chosen an objective function and expressed the government's production constraint above. To complete the formulation of the maximization problem, it remains to add the requirement that the economy be in equilibrium. The conditions that all markets clear can be stated in terms of the vectors x , y , and z .

$$(12) \quad x(q) = y + z$$

The reader may be puzzled that at no place in this formulation has a budget

constraint been introduced for the government. (Other readers may be puzzled by our failure to include only $n-1$ markets in our market clearance equations. These are aspects of the same phenomenon.) Walras' Law implies that if all economic agents satisfy their budget constraints and all markets but one are in equilibrium, then the last market is also in equilibrium. It also implies that when all markets clear and all economic agents but one are on their budget constraints, then the last economic agent is on his budget constraint. In setting up our problem, we have assumed that the household and the private firms are on their budget constraints. Thus, if we assume that all markets clear, this will imply that the government is satisfying its budget constraint,⁷ which we can express as

$$(13) \quad \sum (q_i - p_i)x_i + \sum p_i z_i = 0 \\ = \sum t_i x_i + \sum p_i z_i$$

Alternatively, if we consider the government budget balance as one of the constraints, then it is only necessary to impose market clearance in $n-1$ of the markets.

In this model we can make two price normalizations, one for each price structure. Since both consumer demand and firm supply are homogeneous of degree zero in their respective prices, changing either price level without altering relative prices leaves the equilibrium unchanged. As normalizations let us assume,

$$(14) \quad p_1 = 1, \quad q_1 = 1, \quad t_1 = 0$$

It may seem surprising that it does not matter whether the government can tax a good one. But the reader should remember the budget balance of the consumer. Since there are no lump sum transfers to the

⁷ In an intertemporal interpretation of this model, the government budget is in balance over the horizon of the model, not year by year.

consumer, net expenditures are zero. Thus, levying a tax at a fixed proportional rate on all consumer transactions results in no revenue. (It should be noticed that a positive tax rate applied to a good supplied by the consumer is in effect a subsidy and results in a loss of revenue to the government.)

Welfare Maximization

We can now state the maximization problem. In the statement we shall use the two sets of prices as control variables. It would be a more natural approach to use the taxes which the government actually controls as decision variables. However, once we have determined the optimal p and q vectors we have determined the optimal taxes. Using taxes as decision variables complicates the mathematical formulation and leads to a control problem since the tax vector may not uniquely determine equilibrium.

Rather than calculate the first-order conditions from the formulation spelled out above, we shall alter the problem to simplify the derivation. We have to choose

$$(15) \quad q_2, \dots, q_n, \quad p_2, \dots, p_n, \quad z_1, \dots, z_n$$

to maximize $v(q)$ subject to

$$x_i(q) - y_i - z_i = 0 \quad (i = 1, 2, \dots, n),$$

where y maximizes $\sum p_i y_i$ subject to

$$y_1 = f(y_2, \dots, y_n),$$

and

$$z_1 = g(z_2, \dots, z_n)$$

Since the choice of producer prices can be used to obtain any desired behavior on the part of private producers, we can use any vector y consistent with the production constraint (8). Producer prices are then determined by equation (9). Using the equations

$$y_2 = x_2 - z_2, \dots, y_n = x_n - z_n,$$

we reduce the constraints in (15) to the

single constraint

$$\begin{aligned} x_1(q) &= y_1 + z_1 \\ &= f(x_2 - z_2, \dots, x_n - z_n) + g(z_2, \dots, z_n) \end{aligned}$$

We have therefore simplified the problem (15) to:

$$(16) \quad \text{Choose } q_2, \dots, q_n, \quad z_2, \dots, z_n$$

to maximize $v(q)$ subject to

$$\begin{aligned} x_1(q) - f(x_2(q) - z_2, \dots, x_n(q) - z_n) \\ - g(z_2, \dots, z_n) = 0 \end{aligned}$$

Forming a Lagrangian expression from (16), with multiplier λ ,

$$\begin{aligned} (17) \quad L &= v(q) - \lambda [x_1(q) \\ &\quad - f(x_2 - z_2, \dots, x_n - z_n) \\ &\quad - g(z_2, \dots, z_n)], \end{aligned}$$

we can differentiate with respect to q_k :

$$\begin{aligned} (18) \quad v_k - \lambda \left(\frac{\partial x_1}{\partial q_k} - \sum_{i=2}^n f_i \frac{\partial x_i}{\partial q_k} \right) &= 0 \\ k &= 2, 3, \dots, n \end{aligned}$$

Making use of the equations (9) for producer prices, this can be written

$$\begin{aligned} (19) \quad v_k - \lambda \sum_{i=1}^n p_i \frac{\partial x_i}{\partial q_k} &= 0 \\ k &= 2, 3, \dots, n \end{aligned}$$

Differentiating L with respect to z_k we have

$$(20) \quad \lambda(f_k - g_k) = 0 \quad k = 2, 3, \dots, n$$

Provided that λ is unequal to zero (i.e., that there is a social cost to a marginal need for additional resources), equation (20) implies equal marginal rates of transformation in public and private production and thus aggregate production efficiency as was argued above. The assumption that $\lambda \neq 0$ needs justification. This is provided by the rigorous arguments of Sections III and IV.

If we had introduced several public

production sectors, each described by a constraint like (11), we should have obtained an equation of the form (20) for *each* sector. Thus marginal rates of transformation in all public sectors should be equal, since they are all to be equal to the private marginal rates of transformation. This argument—which we only sketch here, since the conclusion will be proved more directly in the next section—justifies our assumption that there should be production efficiency in the public sector.

Optimal Tax Structure

The relations (19) determine the optimal tax structure, since they show how producer and consumer prices should be related. These equations show that consumer prices should be at a level such that further increases in any price result in an increase in social welfare, v_k , which is the same ratio, λ , to the cost of satisfying the change in demand arising from the price increase. Reintroducing taxes explicitly into the problem we can obtain an alternative interpretation for the first-order conditions.

Since x_i is a function of $p+t$,

$$\frac{\partial x_i}{\partial q_k} = \frac{\partial x_i}{\partial t_k}$$

(p is held constant in this latter derivative.) Consequently, the optimal tax structure, (19), can be rewritten:

$$(21) \quad v_k = \lambda \sum p_i \frac{\partial x_i}{\partial t_k} = \lambda \frac{\partial}{\partial t_k} \sum p_i x_i$$

Since $\sum p_i x_i = \sum q_i x_i - \sum t_i x_i = - \sum t_i x_i$ (by the consumer's budget constraint (6)), we have

$$(22) \quad v_k = - \lambda \frac{\partial}{\partial t_k} (\sum t_i x_i)$$

This last set of equations asserts the

proportionality of the marginal utility of a change in the price of a commodity to the change in tax revenue resulting from a change in the corresponding tax rate, calculated at constant producer prices. Like the first-order conditions for the optimum in standard welfare economics, our first-order conditions are expressions in constant prices. The tax administrator, like the production planner, need not be concerned with the response of prices to government action when looking at the first-order conditions.

If we now make the further assumption that the welfare function is individualistic, we can use equation (5) to replace v_k . The first-order conditions then become

$$(23) \quad x_k = \frac{\lambda}{\alpha} \frac{\partial (\sum t_i x_i)}{\partial t_k}$$

Thus for all commodities the ratio of marginal tax revenue from an increase in the tax on that commodity to the quantity of the commodity is a constant. This form of the first-order conditions has the advantage of showing the information needed to test whether a tax structure is optimal. The amount of information does not seem excessive relative to the data and knowledge which a planner in an advanced country should have.

The statements of the first-order conditions thus far do not directly indicate the size of the tax rates required, nor the impact upon demand that the optimal tax rates would have. In his pioneering study of optimal tax structure, Frank Ramsey manipulated the first-order conditions so as to shed light on the latter question. He employed the concept of demand curves calculated at a constant marginal utility of income. Paul Samuelson reformulated this using the more familiar demand curves calculated at a constant level of utility. We shall return to this question in Diamond-Mirrlees II.

III. Production Efficiency in the Many-Consumer Economy

We have remarked already that many of the results carry over directly to an economy of many consumers, even when lump sum taxation is excluded. We notice at once that the device of expressing welfare as a function of the prices, q , faced by consumers can be used perfectly well. Explicitly, we assume that there are H households, with utility and demand functions u^h and x^h ($h=1, 2, \dots, H$). If, as we may generally suppose, in the absence of externalities from producers to consumers, social welfare can be expressed as a function of the consumption of the various consumers in the economy, $U(x^1, x^2, \dots, x^H)$, it may also be written

$$(24) \quad V(q) = U(x^1(q), x^2(q), \dots, x^H(q)),$$

where we assume that there are no lump sum incomes or transfers that would be influenced by producer prices or government policy. In the case where social welfare depends only on individual utility and there are no externalities, we can write

$$(25) \quad V(q) = W[u^1(x^1(q)), u^2(x^2(q)), \dots, u^H(x^H(q))],$$

where W is presumed to be strictly increasing in each of its arguments.

Using this indirect welfare function, we can carry out the analysis already presented for the one-consumer economy, and conclude in the same way that aggregate production efficiency is desirable. For that argument to be correct, we must confirm that the Lagrange multiplier λ is not zero. Rather than attempt to do this directly, we shall present a different argument for the desirability of production efficiency. A further condition will be required to secure our conclusion. In considering this problem, we shall concentrate on the case where all production is under government con-

trol. The desirability of production efficiency in this case will be seen to imply the same conclusion when there is also a private sector (provided that private producers are price takers, and profits, if any, are transferred to the government). Assume then (as we did in Section I) that all production takes place in the public sector: our problem is to find q that will

$$(26) \quad \text{Maximize } V(q),$$

$$\text{subject to } G(X(q)) \leq 0,$$

where we define $X(q) = \sum_h x^h(q)$ as aggregate demand at prices q . We shall also express the production constraint a little more generally by saying that $X(q)$ is to belong to the production set G , the set of technologically feasible production plans. (Thus the letter G denotes both the production set, and also the function that can be used to describe it; but we shall hardly ever use the *function* G explicitly).

Suppose we establish that, at the optimum for problem (26), production is efficient. Consider an economy with the same technological possibilities, partly under the control of private, competitive producers. The government can induce private firms to produce any efficient net output bundle by suitable choice of producer prices p . In particular, it can obtain the production plan that would be optimal if the government controlled all production. The choice of p does not affect consumer demands or welfare, since pure profit arising from decreasing returns to scale go to the government, and since, any commodity taxes being possible, q can be chosen independently of p . Thus, if the solution to (26) is efficient, the same equilibrium can be achieved when some production is under private control, and is optimal in that case too. Proof that production efficiency is desirable in the "special" case (26) therefore implies that pro-

duction efficiency is desirable in the more general case.

Examples of Inefficiency

Before considering the argument for efficiency, it is useful to consider some limitations on that argument as demonstrated by the following examples of desired inefficiency. It will be recollected that a production plan is efficient if any other feasible production plan provides a smaller net supply of at least one commodity. We shall use a different concept: we say that a production plan is *weakly efficient* if it is on the production frontier. It is possible for a production plan to be weakly efficient without being efficient if the production frontier has vertical or horizontal portions. For matters of economic importance, such as the existence of shadow prices, weak efficiency is all that is required. It is easy to see that if all the prices corresponding to a weakly efficient production plan are positive, the plan is in fact efficient in the usual sense.

Even with this slightly weakened concept of efficiency, it is not necessarily true that, when an optimum exists, optimal production has to be weakly efficient. We present two examples.

Example a is portrayed in Figure 7. It is a one-consumer economy where social preferences, as depicted in the social indifference curve *II*, do not coincide with individual preferences. It is evident that, in the case shown, the optimal production plan is actually in the interior of the production set.

In the second example, social preferences do respect household preferences, but again optimal production lies in the interior of the production set, and is therefore not weakly efficient: suitable producer prices cannot be found, and the social optimum cannot be obtained when there is private control of production.

Example b. There are two commodities and two households. One has utility function x^2y , the other has utility function xy^2 ; each has the nonnegative quadrant $\{(x, y) | x \geq 0, y \geq 0\}$ as consumption set. The first consumer has three units of the first commodity initially; the second, one unit of the second commodity. The welfare function is

$$-\frac{1}{x_1 y_1} - \frac{1}{x_2 y_2}$$

The second commodity can be transformed into the first according to the production relation $x + 10y \leq 0$, ($x \geq 0$). Let the prices of the commodities be q_1, q_2 . Then the first household's net demands are

$$\begin{aligned} & -1 \text{ of the first commodity,} \\ & q_1/q_2 \text{ of the second commodity.} \end{aligned}$$

The second household has net demands

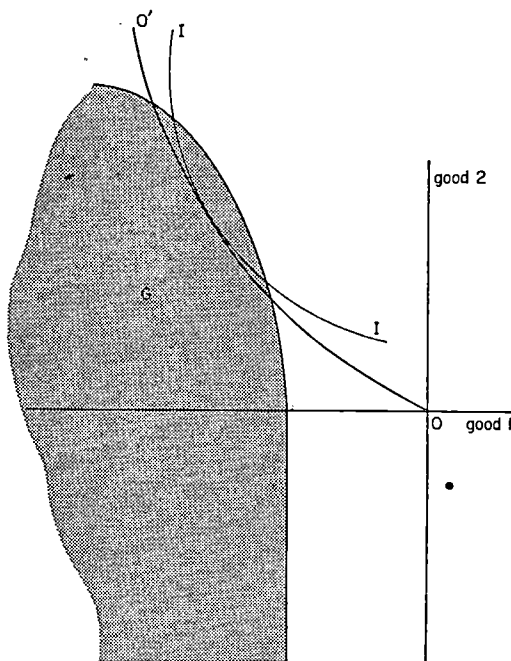


FIGURE 7

$$\frac{1}{3}(q_2/q_1) \text{ and } -\frac{1}{3}$$

Thus, the net market demand for the commodities is

$$x = \frac{1}{3}(q_2/q_1) - 1 \text{ and } y = (q_1/q_2) - \frac{1}{3}$$

These must satisfy

$$x + 10y \leq 0, \quad x \geq 0$$

Welfare is $-q_2/4q_1 - 27q_1/4q_2$ which is maximized when $q_2/q_1 = 3\sqrt{3}$: the corresponding production vector $\sqrt{3}-1, \frac{1}{3}(\sqrt{3}-1)$ is actually interior to the production set, not on the frontier. This example has the unimportant peculiarity that initial endowments of the consumers are on the frontiers of their consumption sets. More complicated examples avoiding the peculiarity have been constructed.

The Efficiency Argument

Despite these examples, the following argument shows that optimal production will generally be on the production frontier. Suppose that the aggregate demand functions, $X(q)$, are continuous. Then any small change in the prices, q , will not change aggregate production requirements by much. Therefore, if optimal production were in the interior of the production set, small changes in consumer prices would still result in technologically feasible aggregate demands. Thus, if we are at the optimum, small changes in consumer prices cannot increase welfare. If we can argue that, at the optimum, there exists a small price change which would increase $V(q)$, we can conclude that production for the optimum must occur on the production frontier. For any unsatiated single consumer,* utility can be increased either by lowering the price of a supplied good or raising the price of a demanded good (as we can see, algebraically, in equation (5)). With a single consumer, we need not argue further, provided the equilibrium involves some trade. When there are many con-

sumers, we can be certain of increasing welfare if we raise some consumer's utility without lowering that of anyone else. If there is a commodity that no consumer purchases, but some consumer supplies (such as certain labour skills); or a good (with positive price) which no consumer supplies, but some consumer purchases (such as electricity), we could alter the price of that commodity in such a way as to bring about an unambiguous increase in welfare. In that case, we conclude that efficient production is required for the maximization of individualistic social welfare. In example *b*, it will be seen that neither of the commodities is supplied, or demanded, by both consumers. The very simplicity of the case appears to be misleading.

A formal presentation of this argument is given in the next section: these technical details can be omitted without loss of continuity. We conclude this section by introducing further taxes into the discussion.

First, consider the case of a poll tax (or subsidy)—that is, a tax is paid by a household on the basis of some unalterable property, such as its sex or age distribution. Such a tax is, of course, a lump sum tax, although its availability is not, in general, sufficient to enable the full optimum to be achieved. To fix ideas, suppose there is a single transfer, τ , to be made to all households. Then welfare can be written $V(q, \tau)$, and we are to

$$(27) \text{ Maximize } V(q, \tau)$$

subject to $X(q, \tau)$ being in G

The standard efficiency argument can be used. Let (q^*, τ^*) be the optimum: if any small change in q or τ would increase V , optimal production, $X(q^*, \tau^*)$ must be on the production frontier (assuming that X is a continuous function). Now a poll subsidy must make everyone better off,

unless some are already satiated, and so must a small increase in subsidy. Thus so long as a poll subsidy is possible (and it surely is) and not every household is satiated, optimal production must be on the frontier.

Adding further tax instruments to the government's armoury in no way weakens the efficiency conclusion. We simply note that if there are other tax variables which are independent of producer prices and quantities, denoted collectively by ξ , we can hold them constant at their optimum values ξ^* , and then apply the efficiency argument to the problem (27) or (26), where V and X are evaluated for $\xi = \xi^*$.

Our final conclusion is that whatever the class of possible tax systems, if all possible commodity taxes are available to the government, then in general, and certainly if a poll subsidy is possible, optimal production is weakly efficient. We would not expect this conclusion to be valid if there were constraints on the possibilities of commodity taxation, or more generally, on the possible relationship between producer prices and consumer demand. The presence of pure profits is one example of such a relationship. To show what goes wrong, suppose, by way of another example, that *no* commodity taxes are possible, but a poll tax is possible, and that part of production is privately controlled, in such a way that it is uniquely determined by producer prices. Then we have to choose a public production vector z and a poll tax τ to

$$(28) \quad \text{Maximize } V(p, \tau)$$

subject to $X(p, \tau) - y(p) = z$ being in G ,

where $y(p)$ is the private production vector when prices are p . Following the argument used above, we consider τ smaller than τ^* , the optimum level, and note that $V(p^*, \tau) > V(p^*, \tau^*)$. This implies that $X(p^*, \tau) - y(p^*)$ is not in G , and therefore z^* , the optimal z , is efficient in G . But the

argument does not imply that the aggregate optimal production plan, $y(p^*) + z^*$ is efficient. Of course, in an economy where all production is under public control, these problems do not arise. Even when some of the q_k are fixed, the efficiency argument holds, for there can be no necessary relation between q and p .

IV. Theorems on Optimal Production

In this section, we explore the existence of the optimum, and the efficiency of optimal production, rigorously. We rely on Debreu (1959) for the results of general equilibrium theory that are required.

Assumptions

There are H households in the economy, each household choosing a preferred net consumption vector x from his consumption set C subject to the budget constraint $q \cdot x \leq 0$ where q is the vector of prices charged to consumers. (Consumption is measured net of initial endowment for convenience, since the latter is unaltered in the analysis.) As usual the net demand vector x has, in general, both positive and negative components corresponding to purchases and sales by the household.

The assumptions used below will be selected from the following list (the superscript h refers to the index of households; all assumptions, when made, hold for all h):

- (a.1) C^h is closed, convex, bounded below by a vector a^h , and contains a vector with every component negative.
- (a.2) The preference ordering is continuous.
- (a.3) The preference ordering is strongly convex. Formally, if x^2 is preferred or indifferent to x^1 and $0 < t < 1$, then $tx^2 + (1-t)x^1$ is strictly preferred to x^1 .
- (a.4) There is no satiation consumption in C^h .

Assumptions (a.1) and (a.2) guarantee the existence of continuous utility functions, which we shall write u^h (see Debreu Section 4.6). Furthermore, under (a.1)–(a.3), when the demand vector $x^h(q)$ is defined, it is uniquely defined. When C^h is bounded, assumptions (a.1)–(a.3) imply that $x^h(q)$ is defined and continuous at all non-zero nonnegative q . (See Debreu, Section 4.10.)

Let us denote aggregate demand by $X(q) = \sum_h x^h(q)$.

It is assumed that all production is controlled by the government. The assumptions on the production possibility set, G , will be taken from the following set:

- (b.1) Every production plan in which nothing is produced in a positive quantity is possible: i.e., if $z \leq 0$, z is in G .
- (b.2) Complete inactivity is possible: i.e., 0 is in G .
- (b.3) G is closed.
- (b.4) There exists a vector \bar{a} such that $z \leq \bar{a}$ for all nonnegative z in the convex closure of G . (i.e., the closure of the convex hull of G).⁸
- (b.5) G is convex.

The welfare function will be denoted by $U(x^1, \dots, x^H)$. When demands are functions of prices only we can define the indirect welfare function as

$$V(q) = U(x^1(q), \dots, x^H(q))$$

Similarly we can define an individual's indirect utility function by

$$v^h(q) = u^h(x^h(q))$$

We shall say that the welfare function *respects household preferences* when U can be written

$$U(x^1, \dots, x^H) = W(u^1(x^1), \dots, u^H(x^H))$$

⁸ When G is convex, this assumption is similar to the assumption that inputs are required to obtain outputs, but permits the government to own a vector of inputs.

with W increasing in each argument. We shall assume

- (c.1) U is a continuous function of (x^1, \dots, x^H)

We can now state our problem as trying to find q^* to maximize $V(q)$ subject to $X(q)$ being in G . A commodity vector will be called *attainable* if it is feasible and if there exists prices such that aggregate demand equals the vector. The set of all such vectors, the *attainable set*, is the intersection of G with the set of vectors $X(q)$ for all nonnegative q .

Existence of an Optimum

If we assume that the attainable set is nonempty and bounded, we obtain

THEOREM 1. *If assumptions (a.1)–(a.3), (b.3), and (c.1) hold, and if the attainable set is nonempty and bounded, an optimum exists.*

PROOF:

Consider an economy in which the consumption sets are truncated by removing from them all points x with $\|x\| > M$, where all vectors in the attainable set satisfy $\|x\| < M$. For this truncated economy, the demand functions are continuous at all price vectors not equal to zero. Since the attainable set, and demands for any q corresponding to an attainable vector, are the same in the original and truncated economies, an optimum for the truncated economy is an optimum for the original economy. In other words, we may, without loss of generality, assume that demands are continuous at $q \neq 0$. Since the demand functions are homogeneous of degree zero in the prices, we can restrict our attention to q satisfying $q \geq 0$ and $\sum_i q_i = 1$.

We next demonstrate that the set $\{q | X(q) \text{ in } G\}$ is closed. Let q_n be a sequence of price vectors converging to q' , with $X(q_n)$ in G for all n . Let x' be a limit point of $\{X(q_n)\}$. Since G is closed, x' is

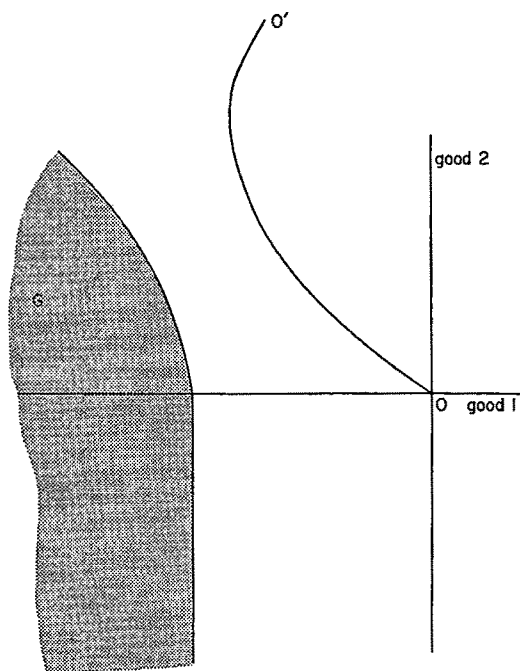


FIGURE 8

in G . At the same time, $x' = X(q')$, by the continuity of X . Thus q' is in $\{q | X(q) \in G\}$, which is therefore closed.

Since the attainable set is nonempty, and prices are in any case bounded, $\{q | X(q) \in G\}$ is closed, bounded, and nonempty. By the continuity of the demand functions, and assumption (c.1), V is a continuous function of q , which therefore attains its maximum on the set $\{q | X(q) \in G\}$.

One criterion for the attainable set to be nonempty follows immediately from the existence of competitive equilibrium in an exchange economy:

THEOREM 2. *If assumptions (a.1)–(a.4) and (b.1) hold, the attainable set is nonempty.*

PROOF:

See Debreu (Section 5.7) for a proof that there exists an equilibrium for the exchange economy with these consumers.

The equilibrium prices result in a feasible demand.

If the production set is taken to be the set of possible production vectors net of government consumption, the assumption that zero production is possible is excessively strong, especially for governments with large military establishments. But it is easy to construct examples of economies not satisfying (b.1) in which there is no attainable point. Consider the one-consumer economy depicted in example c shown in Figure 8.

The boundedness of the attainable set would be implied by the boundedness of the consumption sets, or the boundedness of production, but the following case is more appealing:

THEOREM 3. *If assumptions (a.1) and (b.2)–(b.4) hold, then the attainable set is bounded.⁹*

PROOF:

Suppose the attainable set is not bounded. Then there exists a sequence of attainable vectors x_n such that $\|x_n\|$ is an unbounded increasing sequence of real numbers. There exists an n' such that $\|x_{n'}\| > \|\bar{a}\|$, where \bar{a} is the vector employed in (b.4). Consider the sequence of vectors $(\|x_{n'}\|/\|x_n\|)x_n$ for $n \geq n'$. Each vector is in the convex hull of G (being a convex combination of the origin and x_n). Further the sequence is bounded. Thus there is a limit point, ξ , which is in the convex closure of G and satisfies $\|\xi\| > \|\bar{a}\|$. Let $b = \sum_h a_h$, where a_h are the vectors employed in (a.1). Then $x_n = \sum_h x_n^h \geq \sum_h a_h = b$. Further $(\|x_{n'}\|/\|x_n\|)x_n \geq (\|x_{n'}\|/\|x_n\|)b$. But the latter sequence of vectors converges to zero. Thus $\xi \geq \bar{0}$. This is a contradiction.

⁹ The attainable set will also be bounded if (b.2)–(b.4) hold for the true production set, gross of government consumption, rather than the net production set, G . Thus the assumption that zero production is possible is not of great consequence.

Finally, we should remark that the strong convexity assumption, (a.3), which was made in Theorem 1 can be changed to convexity without affecting the conclusion. All that is required is to replace the continuous functions of the proof by upper semi-continuous correspondences. On the other hand, one can easily construct examples in which an optimum fails to exist because of the absence of continuity.

Efficiency

The following lemma provides two criteria for optimal production to be on the frontier of the production set. It will be used to deduce a theorem about the case where household preferences are respected.

LEMMA 1: *Assume an optimum, q^* , exists. If aggregate demand functions and the indirect welfare function are continuous in the neighborhood of the optimal prices; and if either*

(1) *for some i , V is a strictly increasing function of q_i in the neighborhood of q^* ; or*

(2) *for some i with $q^*_i > 0$, V is a strictly decreasing function of q_i in the neighborhood of q^* ,*

then $X(q^)$ is on the frontier of G .*

PROOF:

Let l_i be the vector with all zero components except the i th, which is one. In case 1, for ϵ sufficiently small $V(q^* + \epsilon l_i) > V(q^*)$. Hence $X(q^* + \epsilon l_i)$ is not in G . Letting ϵ decrease to zero, the continuity of X shows that $X(q^*)$ is a limit of points not in G , and therefore belongs to the boundary of G . In case 2, a similar argument can be made using $V(q^* - \epsilon l_i)$.

These conditions are weak. They are, naturally, independent of production possibilities. It may also be noticed that, when V is a differentiable function of prices, the stated conditions are equivalent to assuming that

(29) It is not the case that $V'(q^*) \leq 0$

Here $V'(q)$ is the vector of first derivatives of V with respect to prices. The equivalence of the conditions of the theorem and (29) is clear if we remember that

$$(30) \quad V'(q) \cdot q = \sum \frac{\partial V}{\partial q_k} q_k = 0,$$

since V is homogeneous of degree zero in q . Therefore $V' \leq 0$ if, and only if, $\partial V / \partial q_k = 0$ when $q_k > 0$ and $\partial V / \partial q_k \leq 0$ in any case.

In the following theorem, we strengthen the assumptions in a different way: they remain notably weak.

THEOREM 4. *If (a.1)–(a.4) and (c.1) hold; if social welfare respects individual preferences; and if either*

(1) *for some i , $x^h_i \leq 0$ for all h , and $x^h_i < 0$ for some h' ; or*

(2) *for some i , with $q_i > 0$, $x^h_i \geq 0$ for all h and $x^h_i > 0$ for some h' ;*

Then if an optimum exists, production for the optimum is on the frontier of the feasible set.

PROOF:

Individual demand functions are continuous in the neighborhood of the optimum and thus aggregate demands and the indirect welfare function are continuous. Since social welfare respects preferences, indirect social welfare can be written as an increasing function of indirect utilities. In case 1, indirect utilities are a nondecreasing function of q_i in the neighborhood of q^* for all h while the indirect utility function of h' is strictly increasing in q_i . Thus V increases with q_i . Case 2 follows similarly.

The assumption of strictly convex preferences made in Theorem 4 is required in the theorem as stated.

Example d: Consider an economy with one consumer whose indifference curves have

different consumers. There are two ways in which we can consider doing this. The country might be geographically partitioned with different consumer prices in different regions. Ignoring migration, and consumers making purchases in neighboring regions, our analysis can be applied to determine taxes region by region. In general the tax structure will vary over the country.

Alternatively, we might consider taxation on all consumer-consumer transactions. Here, too, we would expect to be able to increase social welfare by having these additional tax controls. Neither addition to the available tax structure alters the desirability of production efficiency.

Untaxable Sectors

One problem that arises with a model considering taxation of all transactions is that some transactions may not be taxable, practically or legally. An example of the former might be subsistence agriculture where transactions with consumers are hard to tax while those with firms are not. If the introduction of other taxes (e.g., on land or output) is ruled out, we can accommodate this problem in the model by including subsistence agriculture in the consumer rather than producer sector (or treating it as a second consumer sector). Efficiency would then be desired for the modern and government production sectors taken together; while the tax structure rules would be stated in terms of demand derivatives of the augmented consumer sector rather than of just the true consumers.

Similarly, in an economy without taxes, a public producer subject to a budget constraint is unable to charge different prices to consumers and producers. Lumping together the entire private sector as a single consumer sector, we obtain the conditions for optimal public production of an industry regulated in this manner. This

is the problem considered by Boiteux in the context of costless income redistribution. He also analyzed such an economy with several firms, each limited by a budget constraint.

Foreigners

It is not easy to provide a satisfactory welfare economics for a world of many countries. The study of world welfare maximization is interesting, and, one may hope, "relevant." But it has the serious limitation that its results can seldom be applied to the actions of governments. However altruistic the principles on which a government seeks to act, it has to allow for the actions other governments may take, based on different principles, or for different reasons. (A somewhat analogous problem arises in intertemporal welfare economics.) In the following two subsections, we shall, in order to keep the discussion brief, refer only to the case where the reactions of all other countries are well-defined functions of the actions of the country directly considered. Thus we neglect, reluctantly, those situations that have come to be called "game-theoretic." Also, we shall not consider the problem of formulating a social welfare function in an international setting.

International Trade

So long as we are completely indifferent to the welfare of the rest of the world, and so long as the reactions of other countries are well-defined, international trade simply provides us with additional possibilities for transforming some goods and services into others. The efficiency result then implies that we would want to equate marginal rates of transformation between producing and importing. If there is a monopoly position to be exploited, it should be. If international prices are unaffected by this country's demand, intermediate goods should not be subject to a tariff, but final

good sales direct to consumers should be subject to a tariff equal to the tax on the same sale by a domestic producer.

Sometimes it is not possible to sell goods to foreigners at prices different from those at which they are sold to domestic consumers, although the theory just outlined suggests that foreigners should be treated like producers. As examples, we may cite tourists and commodities covered by special kinds of international agreement. If tourism, say, is an important trading opportunity for the country, and tourists have to be charged the same prices as domestic consumers, this will affect the optimal level of taxes on certain commodities. The general efficiency result is not upset, however. The analysis can be performed by treating tourists as consumers whose income does not affect social welfare.

The authors do not, of course, recommend indifference to the welfare of the rest of the world; although it happens to make the results somewhat neater. International trade provides the country with another set of consumers who can trade with it at prices different from its own consumers: the case (when foreign reactions are well-defined) is similar to the possibility of using different consumer prices in different regions of the same economy. In that case, there is no reason why optimal international trade prices should be the same as producer prices, p , or domestic consumer prices, q .

Migration

In all that has gone before, we have been holding constant the set of consumers in the economy. We can introduce migration in a straightforward manner. Social welfare may be a function of the consumption of every household in the world. Changes in the consumer prices charged in the home country cause migration in one direction or another, and therefore affect wel-

fare in ways we have not previously discussed (such as the effect on the inhabitants of another country of having additional taxpayers join them). But we can still define an indirect welfare function $V(q)$, so long as the reactions of the rest of the world are well-defined. Similarly we can define aggregate demand functions $X(q)$, but these are no longer continuous. For, when a man decides to emigrate, his contribution to aggregate demand changes from x^h to 0.¹⁰ But the number of migrants arising from a small price change may, quite reasonably, be assumed small relative to the population as a whole. We can therefore adequately approximate this situation by considering a continuum of consumers. In this way we can restore continuity to aggregate demand, and to the indirect welfare function. It is to be expected, then, that production efficiency is still desired. Since the derivatives of the demand functions, and possibly also the derivatives of V , will be different when the possibility of migration is allowed for, the optimal tax structure will be changed to reflect the loss of tax revenue when net taxpayers, for example, leave the country. While we do not wish to examine this problem in detail here, we believe that these ideas provide an interesting approach to the analysis.

Consumption Externalities

The schematic view of this problem given above suggests that the basic structure of the results, although not the specific optimal taxes, are unchanged by complications which occur wholly within the consumer sector. Thus, if we introduce consumption externalities that leave aggregate demand continuous we will still obtain production efficiency at the optimum, if we can argue that $V(q)$ has no unconstrained local maximum for finite q .

¹⁰ A similar discontinuity problem arises in the case of tourists' decisions not to visit the country.

The conditions used above are no longer sufficient for this argument since the direct effects of a price change might be offset by the change in the pattern of externalities induced by the price change. Although we have not examined this case in detail, there are a number of cases where arguments similar to those in the no-externality case will be valid.¹¹ Furthermore it seems quite likely to us that efficiency will be desired in realistic settings.

Capital Market Imperfections

While some capital market imperfections affecting firms are complicated to deal with, some imperfections relevant only for consumers can be described as elements solely within the consumer sector. For example, consider the constraint that consumers can lend but not borrow. We must then rewrite consumer utility maximization as subject to a set of budget constraints for the different time periods. In the case of two periods, for example, it would appear as

$$\begin{aligned}
 &\text{Maximize } u(x^1, x^2) \\
 (31) \quad &\text{subject to } q^1 x^1 + s \leq 0 \\
 &\quad \quad \quad q^2 x^2 - s \leq 0 \\
 &\quad \quad \quad s \geq 0
 \end{aligned}$$

where s represents first period savings. From this consumer problem, we still have utility and demand expressible in terms of

prices. We expect that the efficiency result continues to hold. In calculating the optimal formula, though, it becomes necessary to distinguish the time period of the good in question for there are now two Lagrange multipliers giving the marginal utility of income in each of the two periods. For this consumer we have

$$(32) \quad \frac{\partial v}{\partial q_k^1} = -\alpha^1 x_k^1, \quad \frac{\partial v}{\partial q_k^2} = -\alpha^2 x_k^2$$

Since savings are allowed $\alpha^1 \geq \alpha^2$. If the consumer would borrow if he could, $\alpha^1 > \alpha^2$ and the optimal tax structure is altered by this market limitation.

REFERENCES

- W. Baumol and D. Bradford, "Optimal Departures From Marginal Cost Pricing," *Amer. Econ. Rev.*, June 1970, 69, 265-83.
- A. Bergson, "Market Socialism Revisited," *J. Polit. Econ.*, Oct. 1967, 75, 431-49.
- M. Boiteux, "Sur la gestion des monopoles public astreints à l'équilibre budgétaire," *Econometrica*, Jan. 1956, 24, 22-40.
- G. Debreu, "A Classical Tax-Subsidy Problem," *Econometrica*, Jan. 1954, 22, 14-22.
- , *Theory of Value*, New York 1959.
- J. Drèze, "Postwar Contributions of French Economists," *Amer. Econ. Rev. Supp.*, June 1964, 54, 1-64.
- A. Prest and R. Turvey, "Cost-Benefit Analysis: A Survey," *Econ. J.*, Dec. 1965, 75, 683-735.
- F. Ramsey, "A Contribution to the Theory of Taxation," *Econ. J.*, Mar. 1927, 37, 47-61.
- P. Samuelson, "Memorandum for U.S. Treasury, 1951," unpublished.

¹¹ We have benefited from discussions with Elisha Pazner on this subject.

United States Imports and Internal Pressure of Demand: 1948-68

By R. G. GREGORY*

Over the past decade the successive balance of payments crises in the United Kingdom have led to considerable discussions as to the determinants of import behavior. A recurring theme in these discussions is that the demand for imports is related to the pressure on domestic resources. It is often alleged that when the rate of utilization of domestic resources is high, imports increase quite markedly (see, for example, Frank Brechling and J. N. Wolfe, and W. Godley and J. Shepherd). Although there has been very little discussion as to the theoretical link between the level of imports and the internal pressure of demand upon domestic resources, a close reading of the discussion indicates that the contributors appear to believe that the normal income and price responses of traditional theory are not sufficient to explain the cyclical behavior of imports. There is a separate response which arises from excess demand.

The recent gold crises have led to an increased interest in the *U.S.* balance of payments and a similar notion of an excess demand effect has emerged here. When excess demand occurs and capacity con-

straints become operative, it has been argued by W. H. Branson, and W. Lederer, E. Parrish, and S. Pizer, that imports increase quite substantially.

In this paper we look at the *U.S.* imports from the viewpoint of excess demand. We derive a neoclassical demand function for imports in Section I and then in Section II we argue that prices are slow to adjust to their equilibrium levels. Markets are cleared by a number of other variables acting as rationing mechanisms. These variables, which are typically ignored in neoclassical demand analysis, include such factors as the waiting time between the placement and delivery of an order, the alacrity with which suppliers offer trade credit, and even the general enthusiasm of the supplier to seek and obtain new orders. We argue that when there is excess demand in the United States and pressure is exerted upon domestic resources, domestic waiting times increase, credit becomes more difficult to obtain, suppliers are less vigorous in the pursuit of new orders, and consumers therefore turn to foreign suppliers. The result is an increase in imports which would not be predicted from the movement of either relative prices or domestic income. To predict the increase in imports, we need to take account of the extent to which domestic prices deviate from equilibrium as well as the relationship between income and potential output.

In Section III we confront this thesis with *U.S.* quarterly data from 1948-68. We find that there is a substantial effect on

* Senior research fellow at the Institute of Advanced Studies, Australian National University, Canberra. The paper was completed when I was visiting assistant professor at Northwestern University. My thanks to F. Brechling and colleagues of the economics department at Northwestern for encouragement and enthusiasm and to the editor for considerable advice on certain sections of the paper. The project was supported from departmental funds kindly offered to a visitor. Mr Ashley Lyman and Miss F. Wilson provided excellent research assistance. At Canberra, P. J. Lloyd has been very helpful.

imports which can be traced directly to measures of capacity constraints and excess demand. When these variables are included in the demand equation they are very significant, and the explanatory power of the demand equation is considerably increased. In the process the estimates of the conventional price elasticities are increased in magnitude and significance. Over the trade cycle the excess demand effect appears to be capable of generating a variation of 20 percent in the ratio of imports to domestic goods.

I. The Demand Equation

Most import demand equations, such as those of J. Ball and K. Marwah, and R. Rhomberg and L. Boissonneault, are not derived from utility functions and do not have microfoundations. To derive the demand equation from a utility function, as we do in this section, has a certain intuitive appeal and it does provide the step from theoretical concepts to parameter estimates. However, this approach does tend to focus attention upon a number of problems which are difficult to solve: problems of aggregation and questions such as "Whose utility function is it?" We prefer though to face these problems directly rather than specify the demand equation a priori and pretend that the problems do not exist.

The import demand equation that is derived is specified so that the dependent variable is the ratio of the U.S. purchase of foreign goods (imports) to the domestic purchase of goods produced in the United States. This form of demand equation assumes identical income elasticities of demand for both goods (see fn. 2) but this possible disadvantage is more than offset by the advantage to be gained when the equation is fitted to the data. By specifying the demand equation in ratio form the problem of multicollinearity, which usually plagues macro time-series analysis, is re-

duced and we are able to conduct much sharper tests of the numerous hypotheses that are our prime interests in this paper. Furthermore, the number of price terms which enter the demand equation is reduced to two, the price indices of home goods and imports.

We envisage the economy making decisions as to the allocation of expenditure as though it were an entity which possessed a utility function belonging to that class of utility functions which are often called utility trees. The essence of this class of utility functions is that decisions can be thought of as occurring in sequence. For our purposes we hypothesize that decisions occur at two levels. The utility function, therefore, is a two-level utility tree. The first decision is to decide upon the allocation of resources between services and goods. Once this decision has been made (the first level decision), the particular bundle of goods and the particular bundle of services must be chosen (the second level decision).

The first level utility function is assumed to be weakly separable¹ and is written as

$$(1) \quad U = F[\phi_1(X^{(1)}), \phi_2(X^{(2)})]$$

where $X^{(1)}$ and $X^{(2)}$ are vectors of goods and services, respectively. In this paper we are primarily concerned with the second level of the first branch of the utility function; that is, our interest lies in the allocation of resources between goods, and in particular, the allocation between goods produced at home, $X_1^{(1)}$, and goods imported from overseas, $X_2^{(1)}$. The second level utility function which relates these two goods we write as

¹ A utility function is weakly separable with respect to the partition $\phi_1(X^{(1)}), \phi_2(X^{(2)})$ if the marginal rate of substitution $u_i(X)/u_j(X)$ between two commodities i and j from $\phi_s(X^{(s)})$ is independent of the quantity of commodities outside of $\phi_s(X^{(s)})$. (See S. M. Goldman and H. Uzawa.)

$$(2) \quad \phi_1(X^{(1)}) = \left[\sum_{i=1}^2 \delta_i (X_i^{(1)})^{-p} \right]^{-1/p},$$

$$\delta_i > 0, \quad -1 < p = \frac{1-\sigma}{\sigma} < \infty$$

which is the familiar constant elasticity of substitution form.

The problem then is to maximize equation (2) subject to the budget constraint

$$(3) \quad Y^{(1)} = \sum_{i=1}^2 p_i^{(1)} X_i^{(1)}$$

where $p_i^{(1)}$ are prices of $X_i^{(1)}$. $Y^{(1)}$, which is predetermined from the first level allocation problem, is the quantity of resources to be spent on the goods $X_i^{(1)}$. The necessary condition for a solution to this problem is given by the equality of the marginal utility ratio with the ratio of prices, that is,

$$(4) \quad \frac{X_2}{X_1} = \left(\frac{\delta_2}{\delta_1} \right)^{\sigma} \left(\frac{p_1}{p_2} \right)^{\sigma}$$

As we are only concerned with the first branch of the utility function (1), we write the variables of equation (4) and the following expressions without superscripts unless there is some doubt as to their identity.²

² Some of the more obvious properties of this theoretical framework should be mentioned. The advantage, for empirical application, to be gained from writing (1) as weakly separable and (2) as homogeneous is that the second level allocation decision, X_2/X_1 , is independent of the first-level decisions and, as is evidenced by (4), this simplifies the estimation problem considerably. Only two price indices enter (4) and consequently the allocation of resources between the two goods, X_1 and X_2 , is independent of the price of services. The Slutsky income-compensated elasticity of demand for any good, however, does depend upon the first level decisions. This price elasticity, $E_{ii}^{(1)}$ can be written as,

$$E_{ii}^{(1)} = \frac{v_j^{(1)}}{v_i^{(1)} + v_j^{(1)}} \sigma_{ij}^{(1)} - \frac{v_i^{(1)}}{v_i^{(1)} + v_j^{(1)}} E_{ix}^{(2)}$$

where v_j is the expenditure on $X_j^{(1)}$, $E_{ii}^{(1)}$ is the own price elasticity of demand for the i th good and $\sigma_{ij}^{(1)}$ is the elasticity of substitution between the two goods.

As a further refinement we may wish to allow for a change in tastes over the data period. This can be done without altering the basic utility function, or the magnitude of its parameters, if it is assumed that tastes change so as to be X_1 or X_2 augmenting. If the change in tastes is X_1 or X_2 augmenting, it can be entered into the utility function as

$$(5) \quad \phi_1(X^{(1)}) = \left[\sum_{i=1}^2 \delta_i^{(1)} (\eta_i^{(1)} X_i^{(1)})^{-p} \right]^{-1/p}$$

where η_i are the change in taste multipliers. This specification of the effect of a change in tastes will enable us to measure the relative change in tastes between foreign and domestic goods but not the absolute change in satisfaction derived from each individually.³ The notion of a

Since the mean proportion of goods imported over the data period, $v_i^{(1)}/v_i^{(1)} + v_j^{(1)}$, is approximately 13 percent the second term can be overlooked if $E_{ix}^{(2)}$, the elasticity of the i th good with respect to a change in the price of services, is not too large. In these circumstances the elasticity of demand for imports can be approximated by σ_{ij} .

Another important property of this theoretical framework is that, since (2) is homogeneous, the allocation of resources between the two goods, X_1 and X_2 , is independent of the quantity of resources, $Y^{(1)}$, devoted to them. Consequently, this implies $I_i^{(1)} = I^{(1)}$ where I is each respective income elasticity; that is, the income elasticity of each good separately, and the income elasticity of the composite index of goods, are equal. There is no other restriction placed upon the income elasticities. For a further discussion of two-level functions see K. Sato where the concept is applied to production functions.

³ If it is assumed that the multipliers η_i change at a constant exponential rate with time, that is,

$$\frac{\eta_1}{\eta_2} = \eta_0 e^{(r_1 - r_2)t} = \eta_0 e^{rt}$$

where r_i is the change of tastes associated with the i th good, then upon substitution into (6) it will be an easy matter to estimate the relative change of tastes. The coefficient on the time trend in the demand equation will be $(1-\sigma)\lambda(r_1 - r_2)$. If the coefficient on the time trend is estimated to be positive, and $(1-\sigma)$ and λ are also positive (the adjustment specification requires $0 < \lambda < 1$), it follows that r_1 is greater than r_2 and, *ceteris paribus*, the change in tastes has increased the marginal utility of X_2 relative to X_1 . If $(1-\sigma)$ is negative the opposite conclusion would be reached. The in-

change in tastes augmenting the quantity of a good consumed has an analogous counterpart in the theory of production where factor augmenting technological change has been used in a number of studies to measure the sources of output growth, (see, for example, P. A. David and Th. Van De Klundert).

Finally, if we allow for a delayed response arising from costs of adjustment to changes in relative prices on the part of the consumer we can add a Koyck lag⁴ to the demand equation and write it as,

$$(6) \quad \frac{X_2}{X_1} = \left(\frac{\delta_2}{\delta_1} \right)^{\sigma\lambda} \left(\frac{\eta_1}{\eta_2} \right)^{(1-\sigma)\lambda} \cdot \left(\frac{P_1}{P_2} \right)^{\sigma\lambda} \left(\frac{X_2}{X_1} \right)_{t-1}^{(1-\lambda)}$$

This concludes the derivation of the basic demand equation. In the next section the demand equation is considered in more detail in order to show that it may not focus attention upon all the important variables which affect the decision to buy foreign or domestic goods.

II. The Effective Price Relationship

The demand equation derived in the previous section focuses attention upon relative prices, tastes, and a lag in the adjustment to the equilibrium quantities, as the demand variables and relationships which are important in determining the allocation of resources between foreign and domestic goods. However, there is some evidence that actual prices are slow to adjust to their equilibrium values and that in the *short run*, markets are cleared partly

in response to relative waiting times, trade credit terms, rebates, discounts, and the general ability of the sellers to meet customer requirements. For example, in a recent paper, M. D. Steuer, R. J. Ball, and J. R. Eaton (Nov. 1966), have shown that *independent of relative prices*, the relative waiting time between the order and delivery of machine tools is a crucial rationing variable which plays an important part in allocating the flow of new orders among the machine tool industries of the United Kingdom, United States, and West Germany. In another paper (see Ball et al. Sept., 1966), they show that, *independently of relative prices*, capacity constraints affect the level of exports of manufactured goods from the United Kingdom. The implication of both these studies is that markets are being cleared by other variables as well as the usual price variable that is included in demand studies.

If this phenomenon is widespread then conventional economic analysis, which focuses upon income and relative prices as the explanatory variables in a demand equation, may, for particular problems, be quite misleading. In econometric work, for example, the omission of relevant variables from a demand equation can lead to biases and poor predictive performances since, at any given price, different quantities may be demanded depending upon the magnitude of the omitted variables. The same problems may arise with respect to supply curves.⁵

interpretation of a given sign of the coefficient on the time trend therefore depends upon whether σ is greater or less than unity.

⁴ The adjustment model hypothesized is

$$y_t - y_{t-1} = \lambda(y_t^* - y_{t-1})$$

where y_t is $\log(X_2/X_1)$ and y_t^* is the \log of the equilibrium ratio of consumption after consumers have fully adjusted to the prevailing relative prices.

⁵ The notion that empirically there may be no unique short-run demand or supply curves when these are defined in the traditional neoclassical manner is gradually becoming more common. Otto Eckstein and Gary Fromm, for example, in their recent paper in which they discuss the determinants of price, refer to "disequilibrium demand and supply curves" and suggest that disequilibrium rather than equilibrium is the more common situation. If this is an accurate description of markets then most data observations will not lie on either a neoclassical demand or supply curve. If the price is fixed, and there is excess demand, then the gap between the demand and supply curve defines the area

If it is true that in the short run other variables, rather than price, predominantly clear markets, and that disequilibrium situations are more common than equilibrium situations, then there is considerable theorizing yet to be done to specify the dynamic interrelationship between each of the market clearing variables, the relationship between these variables and demand and supply curves, and to analyze in general the economics of disequilibrium situations. Furthermore, for empirical purposes the formulation of demand and supply curves should explicitly take account of disequilibrium situations and introduce each of the market clearing variables explicitly into the functions.

Definition of the Effective Price

Perhaps the simplest modification that can be made to neoclassical theory to meet the criticisms above is to redefine the price variables of equation (6) as *effective prices*. Rather than treat the price of a commodity as a one dimensional variable in the usual textbook manner, it may be more appropriate to define the price as a vector possessing many dimensions. We call this vector the effective price, P_e . Its elements are the actual quoted price P_a , the waiting time, the trade credit terms, rebates and any other ancillary aspects of the contract which are relevant for the decision to purchase or not.⁶ Consideration of *relative effective prices* at home and abroad determines whether the commodity is pur-

chased from domestic or foreign suppliers. From this point onwards we conduct the analysis in terms of the domestic effective price but it should be remembered that a similar analysis applies to foreign effective prices and that it is relative effective prices that should enter the demand equation.

If each of the ancillary aspects of contracts are defined as ω_i the effective price could be written as,

$$(7) P_e = AP_a \prod \omega_i^{\gamma_i} \quad \gamma_i \geq 0, i = 1 \dots N$$

where γ_i are the exponential weights attached to each of the ω_i elements. The equation has been normalized so that the exponent on the actual price term is unity. For elements such as the waiting time the coefficient γ is positive so that an increase in the waiting time between the placement of an order and the delivery of the goods is equivalent to a price increase. Other elements of ω , such as the willingness of the seller to adjust to the requirements of the buyer will possess negative γ exponents. From the buyer's point of view an increase in the willingness of the seller to adjust to his demands is equivalent to a price reduction.

If each of the elements of the effective price are observable, independent of each other, and exogenous to demand decisions, then the effective price ratio could be substituted into (6) for the actual price ratio and the demand equation estimated. However, not all the elements of the vector ω are observable. In some instances, such as the ease at which trade credit is obtainable, the elements are measurable in principle but data series do not exist. In other instances, such as the general enthusiasm and desire on the part of sellers for new orders, the elements are not directly measurable. Consequently, we need to develop proxy variables to measure the elements of the effective price vector. This is done in the next two subsections. First,

in which the price quantity observation must fall. Its exact location will depend upon the extent to which suppliers deviate from the equilibrium supply curve by drawing down inventories and temporarily increasing production. A brief survey of some of the recent literature which expounds similar views in the fields of inflation and employment theory can be found in Edmund S. Phelps.

⁶ An alternative way of thinking about this problem is to follow the approach adopted by K. Lancaster and consider the ancillary aspects of contracts as characteristics of the goods.

we discuss in some detail the derivation of a variable to measure the quoted waiting time between the placement and delivery of an order. This variable will probably be an important component of the effective price. Then a more general theory is developed to relate the elements of ω to observable economic variables. This theory should be regarded as suggestive of possible relationships to enable the development of suitable proxy variables rather than a formally complete model.

Measurement of the Quoted Waiting Time

The waiting time which is relevant for the demand equation is not the actual *ex post* waiting time but the *ex ante* waiting time which is quoted to the prospective customer. The *ex post* waiting time is likely to differ from that quoted if it is uneconomic to calculate the waiting time for each individual customer and if the production flow is subject to some inflexibility. Since these circumstances are likely to exist in most industries the suppliers may form their quotations lq , by using the *ex post* waiting time of the previous period, la , adjusted for its rate of change. In that case, the quoted waiting time in any period might be expressed as,

$$(8) \quad lq = la_{-1} \left(\frac{la_{-1}}{la_{-2}} \right)^x$$

$x > 0$

where x is a positive exponential weight attached to the rate of change of the waiting time and the subscripts refer to the time period. If the waiting time has been increasing, $la_{-1} > la_{-2}$, then suppliers will take account of this fact and quote a waiting time which exceeds that of the previous period. Similarly, if the waiting time has been decreasing the quoted waiting time will be less than that of the previous period.

There are a number of interesting impli-

cations both of theory and of interpretation of the regression equations which might be developed from this specification of the waiting time variable. Ignoring the foreign waiting time for the moment, the domestic quoted waiting time can be included into the demand equation by substituting (8) into (7) and then substituting (7) into (6) to derive,

$$(6') \quad \ln \left(\frac{X_2}{X_1} \right)_t = B' + \lambda \sigma \ln \left(\frac{Pa_1}{Pa_2} \right)_t \\ + \gamma \lambda \sigma (1+x) \ln la_{t-1} \\ - \gamma x \lambda \sigma \ln la_{t-2}$$

where,

$$B' = \ln \left[A^{\lambda \sigma} \left(\frac{\delta_2}{\delta_1} \right)^{\lambda \sigma} \left(\frac{\eta_1}{\eta_2} \right)^{(1-\sigma)\lambda} \left(\frac{X_2}{X_1} \right)_{t-1}^{1-\lambda} \right]$$

and γ is the positive exponential weight that is attached to the quoted waiting time element of the effective price. The empirical results from applying (6') to the data should give a positive coefficient, $\gamma \lambda \sigma \cdot (1+x)$, on the first actual waiting time variable, a negative coefficient, $-\gamma x \lambda \sigma$, on the second waiting time variable and an absolutely larger coefficient on the first waiting time variable than on the second, $|\gamma \lambda \sigma (1+x)| > |-\gamma x \lambda \sigma|$. Furthermore, by taking the estimate of $\lambda \sigma$ from the coefficient on the relative actual price term the parameters γ and x , which should both be positive, can be identified.

Data on the *ex post* waiting time for goods produced in the United States are not available but the waiting time can be approximated⁷ by dividing the level of unfilled orders, uo , by the current production rate, Q , so that

$$(9) \quad la \simeq \frac{uo}{Q}$$

⁷ If there is little variance in the rates of growth of production and new orders then this approximation is reasonably accurate. The degree of accuracy worsens if either series fluctuates.

The substitution of (9) into (8) completes the derivation of lq which is an estimate of one of the elements of the vector ω .

Since the above technique to estimate the quoted waiting time may be inaccurate and/or the relationship between the various elements of the ω vector may be too complex for their variation to be approximated by the quoted waiting time, we now attempt to derive a number of proxy variables from a more general theory of the relationship between the elements of the effective price vector and observable economic variables.

A Theory of the Effective Price Relationship

First we assume that firms are imperfect competitors in the product market in the sense that they can control P_a as well as all the elements of ω . They supply only the domestic market.

Second, to develop the theory as simply as possible we assume that the capital stock of each firm is fixed. We do, however, allow for variations in the inventory level of the firm.

Third, we define ω^* to be the long-run equilibrium value of the vector ω . In this paper our primary interest is in disequilibrium situations so ω^* will be regarded as a constant.⁸ Likewise I^* is the long-run equilibrium level of inventories which, at this stage, is also regarded as a constant.

Fourth, there is a cost function defined upon different levels of ω such that the further ω departs from ω^* the greater the cost the firm must bear. If ω exceeds ω^* and waiting times are too long and credit terms too stringent then customer good-

will and future orders will be lost. If ω is less than ω^* and waiting times are too short and credit terms too cheap then production planning may be more difficult and the amount of capital required to finance the trade credit more than optimal. The cost of deviations of ω from ω^* , however, is always less than that of not clearing the market. We would argue that the loss of goodwill that occurs when a firm simply refuses to supply the product at the current price and current waiting time exceeds the loss of goodwill which may occur if the waiting time or other elements of ω are increased. Similarly, there is a cost function defined upon different levels of I such that the further I departs from I^* , the greater the cost the firm must bear.

From these assumptions the demand, Q_D , for the product of the typical firm can be written as,

$$(10) \quad Q_D = g(P_a, \omega)$$

and its supply curve⁹ defined as

$$(11) \quad Q_S = h(\bar{K}, T, W, g(P_a, \omega), \omega, I)$$

where \bar{K} is the fixed quantity of capital, T the level of technology, W the wage rate of labor, and I the inventory level.

Fifth, we assume that firms are reluctant to change P_a frequently in the short run. Their reluctance may be rationalized by postulating that the administrative costs of changing the price are considerable and that frequent price changes destroy the goodwill of customers. In large companies with many products, many outlets, and a large sales force it may even be physically impossible to change prices in immediate response to changes in demand and supply conditions.

Sixth, since firms are reluctant to change P_a in the short run, markets will be cleared

⁸ In a long-run context, ω^* might be an endogeneous variable. There is some evidence that some of the elements of ω^* differ across countries. Steuer, Ball, and Eaton (Nov., 1966) discovered that the United Kingdom is known as a "low price-long delivery lag" supplier of machine tools in contrast to the United States which is a "high price-short delivery lag" supplier.

⁹ The supply equation (11) states that the quantity supplied which will maximize profits can be calculated from the demand curve and those factors which determine marginal cost.

by other devices so that any demand at current effective prices will be satisfied. Consider, for example, a situation in which equilibrium is disturbed by a once and for all injection of excess demand, Z . Excess demand either positive or negative, is defined to exist when

$$(12) \quad Z = Q_D^* - Q_S^* = g(Pa, \omega^*) \\ -h(\bar{K}, T, W, g(Pa, \omega^*), \omega^*, I^*)$$

is either positive or negative. This definition of excess demand is not quite parallel to the usual definition because here excess demand is measured as the gap between the quantity demanded and supplied at current prices but *equilibrium* values of ω and I . When equilibrium is disturbed by a once and for all injection of excess demand, it is hypothesized that the responses of the firm can be divided into three separate groups and that these groups of responses follow each other sequentially through time. Within each group, however, it is assumed that the responses occur simultaneously. These sets of decisions consist first of an inventory response, then an ω , output, and a further inventory response, and finally a price response. We now turn to a further elaboration of this behavior pattern. Not all the disequilibrium relationships will be developed, only those which enable ω to be measured by variables for which there are existing data series.

The first response to a once and for all injection of excess demand, as defined above, is that the inventory level changes to meet the gap between the production rate and the quantity demanded. This response can be written as,

$$(13) \quad \frac{I_{-1}}{I^*} = j(Z_{-1}); \text{ where } \frac{I_{-1}}{I^*} \geq 1 \text{ as } Z_{-1} \leq 0$$

I^* is the equilibrium level of inventories which is assumed to be a constant, and the injection of excess demand initially occurs in period $t-1$. There are a number of

reasons why this might be the first response. In most firms there will be both a lag between the advent of excess demand and its recognition and lag between that recognition and action. During this time the inventory level will alter as a result of the excess demand conditions. Furthermore, in most firms the costs of adjustment attached to the change in the inventory level may be small compared to the costs of adjustment attached to the other economic variables. If the excess demand persists the inventory level continues to fall and the costs imposed upon the firm increase. It then becomes optimal to adjust some of the other variables.

The second set of responses consist of three actions which are mutually determined. These involve a further change in the inventory level, an output response, and a change in the effective price by altering all or some of the elements of ω . The functional relationship which defines the combinations of these responses which will meet the excess demand can be written as,

$$(14) \quad k\left(\frac{I}{I_{-1}}, \frac{\omega}{\omega^*}, \frac{Q}{Q_{-1}}; Z_{-1}\right) = 0$$

where Q is the rate of output. Equation (14) describes a behavioral relationship between these three endogenous variables and excess demand. For any given Z_{-1} , equation (14) can be thought of as a three dimensional surface, concave to the origin, where for ease of exposition the axes are defined as the ω response, the output response, and the *inverse* of the inventory response. This surface connects together all those points which would ensure that demand is satisfied. For movements around this surface the partial derivatives of each of the ω , output and *inverse* of the inventory responses with respect to each other and keeping Z_{-1} fixed are all negative. For higher levels of Z_{-1} the surface moves outwards from the origin.

The combination of these three re-

sponses is chosen by minimizing, subject to (14), the sum of the cost function defined upon the level of output and the cost functions defined upon the deviations of ω and I from their equilibrium levels. It is assumed that (14) and the sum of the cost functions are such that the expansion path, for different levels of Z_{-1} , is monotonic; that is, larger increases in output and ω are always associated with greater reductions in inventories. This expansion path can be written as,

$$(15) \quad \omega = m \left(\frac{I}{I_{-1}}, \frac{Q}{Q_{-1}} \right)$$

where the association between changes in ω and changes in inventories is negative and between changes in ω and changes in output positive. Since we are concerned with short-run responses, ω^* and I^* are regarded as constants and omitted from the functional specifications. This completes the description of the second set of responses.

Since the equilibrium values of I and ω are assumed to be constant the level of excess demand, as defined above, and hence long-run disequilibrium, will persist despite this second set of responses. This situation leads to the third set of responses. The third response is to change the price level, thereby remove excess demand, and to bring the ω elements, inventories and the output rate to long-run equilibrium levels. For our purposes, only the price response is needed and it can be written as,

$$(16) \quad \frac{P_{+1}}{P} = n(Z_{-1}),$$

where

$$\frac{P_{+1}}{P} \geq 1 \text{ as } Z_{-1} \geq 0$$

This expression, which has a long history in discussions of excess demand (see, for

example, Paul Samuelson), completes the model.

We are now in a position to derive proxy variables for the vector ω which can be substituted into the effective price relationship (7) which in turn can be substituted into the demand equation (6). Equation (15) provides one set of proxy variables and if the theory outlined above were correct then this is all that is needed. However, given the degree of complexity of the dynamic interrelationships of the various responses perhaps the best strategy, at this stage, is to derive other sets of proxy variables to provide further tests of the theoretical framework. This can be done in two different ways. The first and simplest is as follows. From (15) it is known that ω and excess demand Z_{-1} are positively associated. From (13) and (16) it is known that inventory reductions in period $t-1$ and price increases in period $t+1$ are also positively associated with excess demand. Therefore, ω could be measured in terms of these variables,

$$(13') \quad \omega = j'(I_{-1})$$

or

$$(16') \quad \omega = n' \left(\frac{P_{+1}}{P} \right)$$

The monotonicity of the original relationship will give unambiguous sign associations between these proxy variables and ω . There is a positive association between the ω elements of the effective price and the future rate of change of prices and a negative association between the ω elements of the effective price and the past change of inventories from their normal level.

The second and more complex procedure is to use (14) to substitute for any of the variables on the right-hand side of (15) (thus introducing Z_{-1} explicitly into (15)) and then use (13) or (16) to substitute for Z_{-1} to introduce into the range of proxy

variables price and inventory changes as well as one of the original variables on the right-hand side of (15). It is fairly easily shown, however, that this technique does introduce some ambiguity in the sign of the coefficient of the original variable which appeared in (15).

As an example of this procedure we can use (14) to derive Q/Q_{-1} in terms of I/I_{-1} , ω and Z_{-1} and then substitute this relationship for Q/Q_{-1} into (15). Then for Z_{-1} we can substitute either the inverse of (13) or (16), and rearrange the expression to give, respectively:

$$(15') \quad \omega = m' \left(\overset{?}{\frac{I}{I_{-1}}}, \bar{I}_{-1} \right)$$

or

$$(15'') \quad \omega = m'' \left(\overset{?}{\frac{I}{I_{-1}}}, \overset{+}{\frac{P}{P}} \right)$$

where the symbols above the variables define the predicted sign pattern of their coefficients.¹⁰ Although the sign of particular coefficients may be ambiguous the theory often predicts relationships between them. For example, if (15') is specified as a log-linear function it becomes

$$(17) \quad \log \omega = \alpha_1 \log \frac{I}{I_{-1}} + \alpha_2 \log I_{-1}$$

¹⁰ As these substitutions are a little complicated it may be useful to spell the procedure out in a little more detail. Consider the derivation of (15'). We can rearrange (14) to write

$$(1') \quad \frac{Q}{Q_{-1}} = r \left(Z_{-1}, \omega, \frac{I}{I_{-1}} \right) \quad r_1, r_3 > 0, r_2 < 0$$

Substituting (1') into (15) gives

$$(2') \quad \omega = s \left(\overset{+}{\frac{I}{I_{-1}}}, Z_{-1}, \overset{-}{\omega}, \overset{+}{\frac{I}{I_{-1}}} \right)$$

where the associations between the variables on the right-hand side and ω are given above the variable. The important fact to notice is that the first and fourth term of (2'), I/I_{-1} , possess different signs. Therefore, inverting (13), substituting for Z_{-1} in (2') and then rearranging we derive (15') in the text.

or rearranging

$$(17') \quad \log \omega = \alpha_1 \log I + (\alpha_2 - \alpha_1) \log I_{-1}$$

From (15') above it is known that α_1 is of ambiguous sign but that α_2 is negative. Consequently if α_1 is positive then $|\alpha_1| < |\alpha_2 - \alpha_1|$. If α_1 is negative then $(\alpha_2 - \alpha_1)$ cannot be greater than $-\alpha_1$ without violating the restriction that α_2 be negative. Likewise, if (15'') is specified as a log-linear relationship and the rate of change of inventories entered in the regressions as the inventory level and its lagged value, then the signs of the coefficients should alternate although there is no restriction on whether the first coefficient is positive or negative.

To conclude this section, the relationship between the effective price and excess demand on the one hand, and the inter-relationship between the elements of the effective price on the other, are likely to be more complex than the simple theoretical relationships sketched above. It is conceivable that each measure of the ω vector that we have discussed will capture a different aspect of these relationships. In the next section therefore we will experiment with different combinations of these measures.

III. The Empirical Results

In the first part of this section the model is used to explain the quarterly variation of the imports of general merchandise into the United States from 1948 to 1968. Total imports, of course, contain a significant proportion of raw materials and producer goods,¹¹ both of which may not at first glance seem to fit easily into the demand framework developed in Section I. However, I feel justified in choosing this level

¹¹ If we were primarily interested in raw materials or producer goods the utility functions of Section I would be replaced on CES production functions. There is no a priori reason why production functions should not consist of more than one level. The article by Sato contains a full discussion of two-level production functions.

of aggregation because it is of direct policy interest and, more importantly, the general notion of an effective price is applicable to *all* goods that are obtainable from both foreign and domestic suppliers. In a later part of this section, I briefly examine more disaggregated data. I also consider a simpler model which measures capacity constraints directly.

The Imports of General Merchandise

This section presents the results of a number of regressions which include different proxy variables to measure the effective price. It is found that, independently of the choice of the particular set of proxy variables suggested in Section II, the concept of a domestic effective price proves to have considerable ability to explain the behavior of the ratio of U.S. imports to the domestic production of goods. All coefficients possess the correct sign and almost all are significant at conventional levels. Insignificant coefficients are occasionally encountered when the rate of production is used as one of the proxy variables to measure the ω elements.¹² These results are not reported. All other coefficients of proxy variables are very significant.

The only failure of the model is associated with the measurement of the foreign effective price. As a result of data difficulties the proxy variables used were very different from those which were used to measure the domestic effective price. At this time I have had no success with any of these variables and they are excluded from the published results.¹³

¹² The referee has pointed out that among the third set of responses, which have not been developed in Section II, there may be a further rise in output accompanying the fall in ω . Consequently unless the sets of responses developed in Section II coincide with our data time periods the variable Q/Q_{-1} may well be a bad proxy variable for ω .

¹³ At this juncture this should not be too serious a source of concern. When all foreign countries which

Table 1 lists a selection of the results. In all regressions the R^2 are between .92 and .95. The first equation of Table 1 is the standard import function (6) where the price term is expressed as the relative actual¹⁴ price rather than the effective price ratio. This equation does not perform as well as those that include the effective price as the independent variable. Furthermore when the effective price is included as an independent variable there are a number of common features running through each of the regressions.

First, the short-run elasticity of substitution, with respect to changes in the actual price ratio, increases quite considerably although the new estimate depends upon which specification of the effective price is used. The short-run elasticity of substitution, which is approximately equal to the own price elasticity of demand for imports (see fn. 2), is usually less than unity but the long-run price elasticity always exceeds unity. The typical value of the long-run elasticity is in the vicinity of three which is a much larger estimate than has been found in earlier studies (see, for example, Ball and Marwah).

Second, the significance level of the actual price elasticity increases quite

supply the United States with imports are grouped together then, at any point of time, some will be experiencing positive excess demand and others negative excess demand so that their aggregation will involve some cancelling. Furthermore, the proxy variables used are, on a priori grounds, not as satisfactory as their U.S. counterparts. I have tried the rate of change of import prices, the deviation of world imports around their trend and the deviation of European production around its trend.

¹⁴ As a price index for imports I used either the implicit price deflator of imports or a unit value index. Both were taken from the *Survey of Current Business* and are subject to a number of special difficulties, see, for example, the Joint Economic Committee *Hearings* 1961. The implicit price deflator usually gave slightly higher estimates of the elasticity of substitution. The price index of imports used in Table 1 is the implicit price deflator of imports.

TABLE 1—THE RATIO OF IMPORTS TO DOMESTIC GOODS PRODUCTION REGRESSED
AGAINST VARIOUS SPECIFICATIONS OF THE EFFECTIVE PRICE RATIO
Quarterly Data 1948-68 (*t*-values in parentheses)

Equations	1	2	3	4	5	6
Constant	.4800 (2.85)	.8528 (5.13)	1.0683 (6.31)	.8731 (2.90)	1.2316 (4.21)	.7149 (4.19)
$\frac{P_1}{P_2}$.2308 (2.10)	.8906 (4.68)	1.1551 (5.85)	.8478 (4.55)	1.1412 (6.02)	.4311 (3.68)
<i>t</i>	.0009 (1.86)	.0012 (2.63)	.0011 (2.56)	.0010 (2.30)	.0010 (2.32)	.0010 (2.21)
$\frac{X_{2t-1}}{X_{1t-1}}$.8075 (11.54)	.6889 (10.44)	.6097 (9.20)	.7180 (10.40)	.6232 (9.15)	.7143 (10.15)
l_{at}		.5712 (5.00)	.4764 (4.26)	.5213 (4.22)	.3989 (3.37)	
l_{at-1}		-.4340 (4.51)	-.3246 (3.38)	-.3747 (3.28)	-.2288 (2.06)	
$\frac{P_{1t}}{P_{1t-1}}$			1.3971 (2.26)		1.3353 (2.25)	1.9064 (3.01)
$\frac{P_{1t-1}}{P_{1t-2}}$			1.7333 (2.49)		2.0104 (2.97)	1.2780 (1.82)
$\left(\frac{I}{X_1}\right)_{t-1}$				-.6479 (2.32)	-.7559 (2.94)	
$\left(\frac{I}{X_1}\right)_{t-2}$.5789 (2.28)	.6084 (2.62)	
S.E.R.	.048	.042	.039	.040	.038	.044
R^2	.92	.94	.95	.94	.95	.93
<i>D</i>	1.34	1.25	.68	1.87	1.21	.23

Symbols: X_1 =the domestic production of goods, constant prices; X_2 =the value of general imports, constant prices; P_1 =implicit price deflator of domestic goods output; P_2 =implicit price deflator of imports; l_a =average delivery lag; *t*=time; *I*=manufacturing and trade inventories.

Source: Data taken from the *Survey of Current Business*; *D* is a new statistic devised by J. Durbin to test for serial correlation when the equation contains lagged dependent variables. All variables are expressed as logarithms, seasonally adjusted, and measured in constant prices.

considerably¹⁵ as proxy variables for ω are added to equation (6). In almost all instances the level of significance is doubled.

¹⁵ Both these results are to be expected. If the demand equation is misspecified to include only the actual price as a measure of the effective price variable then this misspecification is similar to an errors in variable problem. The omitted deviations of ω around ω^* are very similar to a random measurement error. Under these

Third, the coefficient on the time trend is not very sensitive to the way in which the effective price is measured. Since the coefficient on the time trend is positive,

circumstances, as elements of the effective price are added to the regression equation the price variable is gradually purged of its errors and the downward bias is removed. Furthermore, the significance level should increase as the actual price variable is allowed to capture more and more of its true effect.

and the long-run elasticity of substitution is always greater than unity, the time trend coefficient implies that tastes have been changing towards foreign goods (see fn. 3). The effect of the implied bias of tastes towards foreign goods, however, is always small, being somewhere between $-.03$ and $-.002$ percent per quarter.¹⁶

From the preceding paragraphs it is clear that the substitution of the effective price for the actual price has led to a significant improvement in the demand equation. We now turn to an analysis of the various specifications of the effective price.

Consider the waiting time variables first. In Section II, it was suggested that the quoted waiting time would depend upon the actual waiting time and its rate of change. The implications of this hypothesis (the reader should refer back to (6')) are, a positive coefficient on the waiting time variable, a negative coefficient on the waiting time variable lagged one quarter, and that the absolute value of the positive coefficient should exceed that of the negative coefficient. The results of each equation in Table 1 satisfy these restrictions. As further variables are added to the regression equation 2, the only discernible effects upon the waiting time coefficients are a slight reduction in their absolute magnitude and significance. The coefficients always remain significant at conventional levels. In the regression equations we found that the current waiting time and the waiting time lagged one quarter (see (6')) performed better than the waiting time lagged one and two quarters, respectively. This should only

give rise to some concern if there is some a priori reason why the theoretical time periods of Section II should exactly coincide with quarters.

As a further step we can decompose the estimated coefficients on the waiting time variables to derive estimates of the exponent, γ , and the weight, α , which is attached to the rate of change effect¹⁷ (see equations (7) and (8)). We find that γ is approximately .15. This implies, for example, that a 50 percent increase in the quoted waiting time is equivalent to an 8 percent increase in the actual price.¹⁸ The estimates of α lie between 2.5 and 3.0. These estimates imply, for example, that a 20 percent change in the actual waiting time from one quarter to the next gives rise to an 80 percent increase in the quoted waiting time.

The third regression equation includes the quoted waiting time, (see equation (8)) and as another measure of the ω elements of the effective price, the rate of change of the implicit home goods deflator, equation (16'). The theory of Section II predicts a positive coefficient on the rate of change of price variable. In the regressions reported, the rate of change of prices for two quarters has been included. That both are positive and significant might be explained by the aggregation of different industries

¹⁷ As an example of these derivations consider regression equation (4) where $\lambda\sigma \approx .84$, $\gamma\lambda\sigma(1+\alpha) \approx .52$, $-\gamma\lambda\sigma\alpha \approx -.37$. Consequently, $\gamma\lambda\sigma \approx .52 - .37$ and $\gamma \approx \frac{.15}{.84} \approx .17$. To derive α we have $\gamma\lambda\sigma\alpha \approx .37$, and therefore $\alpha \approx \frac{.37}{.15} \approx 2.6$.

¹⁸ Taking the differential of (7) with respect to P_a and l_q gives,

$$dPe = \gamma Pe \frac{dl_q}{l_q} + Pe \frac{dP_a}{P_a}$$

Setting dPe equal to zero, rearranging and substituting the estimate of γ from equation (2), (see fn. 17) gives

$$\frac{dP_a}{P_a} = -.17 \frac{dl_q}{l_q}$$

and therefore a 50 percent increase in the quoted waiting time is equivalent to an 8.5 percent increase in the actual price.

¹⁶ The coefficient on the time trend is $\lambda(1-\sigma)r$. If we use equation (1) as an example we can solve for r as follows: $\lambda \approx .2$, $\sigma \approx 1.15$ and $\therefore -.2(.15)r = .0009$ which gives r equal to $-.03$ percent per quarter which means that $|r_2|$ exceeds $|r_1|$ and tastes have changed towards foreign goods. If, however, there is a time trend in the inventory-output ratio then the coefficient in the time trend would include this effect and r can no longer be identified.

in which the lags between the relationships developed in Section II differ. In conclusion, the greater the rate of change of prices over the last two quarters, the greater the value of the ω elements during and before those quarters; and consequently, given the lag between the placement of an import order and its arrival, the greater the value of current imports.

The fourth equation includes the level of inventories and its rate of change. Inventories have been normalized by dividing through by the production rate. The coefficients possess the correct sign pattern (see the discussion of equation (17') in Section II) and are both significant. The condition derived in Section II that if α_1 , the first coefficient, is negative then $(\alpha_2 - \alpha_1)$, the second coefficient, should not be greater than $-\alpha_1$, is also satisfied.

The fifth equation includes the quoted waiting time and the variables of equation (15'') as a measure of the ω elements. All coefficients again satisfy the restrictions of Section II and are at the same time significant at conventional levels. The coefficients on the inventory variables should be approximately of equal magnitude but opposite in sign and the coefficients on the rate of change of prices positive.

The sixth equation omits the quoted waiting time and measures the ω elements by an approximation to equation (16'). Again the coefficients possess the correct sign and are significant. We now turn to a discussion of the adjustment speeds implied by these equations.

Each of the equations in Table 1 contains a lagged dependent variable. The advantage to be gained from this specification is that it enables an estimate to be made of the lagged response of consumers to changes in the effective price ratio. The estimates of the coefficients on the lagged dependent variable imply an average lag of one and a half to four quarters and a long-run price elasticity, which is two to

three times larger than the short-run price elasticity. We also find that as the specification of the demand equation improves, that is as proxy variables are included to measure the effective price, then the estimate of the response speed is increased. The average lag is reduced from four quarters (equation (1)) to approximately one and a half (equations (3) and (5)). The use of lagged dependent variables, however, does introduce a number of problems which may give rise to some concern.

It has been shown by E. Malinvaud and others that when there is both serial correlation amongst the error terms and a lagged dependent variable then the application of least squares regression will bias the coefficient on the lagged dependent variable either up or down depending on whether the serial correlation is negative or positive. The estimate of the adjustment speed and the long- and short-run multipliers will therefore be biased. Furthermore, the estimate of the variance of the coefficient on the lagged dependent variable will also be biased but the evidence that is available suggests that this bias will be small (see, for example, Malinvaud). It is therefore of some interest to know whether serial correlation is present in the equations of Table 1.¹⁹

¹⁹ A somewhat related problem to this is that the inclusion of lagged dependent variables may well give rise to misspecification problems. Z. Griliches has pointed out that if the correct specification of a model includes serial correlation of the error term but excludes the lagged dependent variable then the inclusion of a lagged dependent variable and the application of least squares will usually result in an improved fit to the equation and a significant estimate of the coefficient on the lagged dependent variable. Consequently, the results of including a lagged dependent variable may well be interpreted as supporting a partial adjustment model when in fact this is not the correct specification. As an example of this problem consider the following simple model. The equations

$$(1) \quad \begin{aligned} y_t &= a + bz_t + u_t \\ u_t &= \rho u_{t-1} + \epsilon_t \quad \epsilon_t \sim N(0, \sigma^2) \end{aligned}$$

could be written as

$$(2) \quad y_t = a(1 - \rho) + bz_t - \rho bz_{t-1} + \rho y_{t-1} + \epsilon_t \quad (\text{over})$$

There are two reasons why we think that it is unlikely that the error terms of these equations are serially correlated. First, the regressions were run again to include the dependent variable lagged two quarters. If serial correlation were present it is more than likely that the coefficient on the dependent variable lagged two periods would be significant (see fn. 19). In all instances this coefficient was insignificantly different from zero and the other coefficients were not significantly changed. Second, we have applied a test devised by J. Durbin which also suggests that there is no first-order serial correlation. This sta-

tistic, D , is asymptotically valid and will test for first-order positive serial correlation in the presence of lagged dependent variables. The statistic is defined as

$$(18) \quad D = a \sqrt{\frac{n}{1 - n\hat{\sigma}(b)}}$$

where $a = 1 - \frac{1}{2}d$, d is the Durbin-Watson statistic, n is the sample size, and $\hat{\sigma}(b)$ the estimate of the variance of the coefficient of the lagged dependent variable. The statistic D is tested as a standard normal deviate so that if D is positive and greater than 1.96 then the null hypothesis is rejected and it is concluded that positive first-order serial correlation exists. It can be seen from Table 1 that D is uniformly less than 1.96 and therefore we conclude that there is no positive first-order serial correlation and therefore no serious bias with respect to the estimate of the adjustment speed.

It would appear from the results, and the discussion so far, that the concept of an effective price as a rationing device is a very powerful one. It is very successful at explaining the decisions of large groups of people to buy foreign or domestic goods. In order to get some feel for the quantitative importance of the variations of the ω elements of the effective price vector we conducted the following experiment. The relative actual price and time trend variables were set at their mean values and the observations on each of the excess demand variables were placed in the regression equations and the predicted X_2/X_1 ratios calculated. In all instances the Koyck lag was allowed to remain operative so that the predicted values are to be interpreted as short-run responses. The proportionate range of the predicted foreign-domestic goods ratio under different specifications of ω proved to be approximately 20 percent. Consequently, it would appear that the notion of an effec-

which, apart from the different error specification, is operationally equivalent to adding a partial adjustment response to the model

$$(3') \quad y_t = a + bz_t - cz_{t-1} + v_t$$

to give

$$(4') \quad y_t = \lambda a + \lambda bz_t - \lambda cz_{t-1} + (1 - \lambda)y_{t-1} + \lambda v_t$$

Griliches has suggested that these two models (2') and (4') may be distinguished by the following rule of thumb. If the coefficient on z_{t-1} is equal to minus the product of the coefficient of z_t and y_{t-1} then this can be taken as evidence that the model is not a partial adjustment model. If (2') and (4') are generalized to include more exogenous variables then it can be seen that the general form of these equations is very similar to the general form of the equations listed in Table 1. It is quite conceivable therefore that what we have interpreted to be a partial adjustment model is, in fact, the outcome of positive first-order serial correlation. For example, the coefficient on the lagged waiting time variable in regression 2 is not that different from minus the product of the coefficient on the current waiting time and the coefficient on the lagged dependent variable, that is $-(.57)(.68) = -.39 \approx -.43$. Consequently our interpretation of the lagged dependent variable may be incorrect. There is, however, other evidence which supports the partial adjustment hypothesis.

A generalization of (2') indicates that all exogenous variables, lagged one period, should enter into the estimated equation. In the context of the regressions reported in Table 1 this would suggest that both the lagged values of relative prices and the rate of change of prices should enter into the regressions with negative and significant coefficients. It can be seen from Table 1 that the restriction on the coefficient of the lagged rate of change of prices does not hold. Furthermore, in a number of unreported regressions it was found that when the lagged relative price term was included its coefficient was insignificantly different from zero.

tive price, which in the short run clears the market to a large extent by variations in its non-price component, is quantitatively an important concept.

The Rate of Capacity Utilization

In many of the discussions of the relationship between imports and the pressure of demand upon domestic resources, the rate of capacity utilization in the manufacturing sector is often explicitly mentioned as a causal factor explaining some of the variation in imports over the business cycle (see, for example, Branson). In this paper we have argued that it is not the capacity constraints per se that affect the allocation between domestic and foreign goods but rather the relative effective prices that face consumers. However, if the effective domestic price varied directly with the rate of capacity utilization of the domestic manufacturing sector then the rate of capacity utilization may serve as a simple proxy variable for the cyclical variations of the effective price.

In Table 2 we list some of the results obtained from hypothesizing that ω is a log-linear function of the rate of capacity utilization in the manufacturing sector. The capacity variable was subject to a number of transformations not all of which appear in Table 2. In different regressions it was lagged one, two, or three quarters and in order to test for non-linearities the variable was subdivided into observations when the capacity utilization rate was less than 81 percent (c_1), between 81 and 89 percent (c_2), and greater than 89 percent (c_3). We also tested the relationship between ω and the rate of capacity utilization for cyclical asymmetries by separating the capacity variable into one series of observations when the rate of capacity utilization was increasing ($c \uparrow$) and another series of observations when the rate of capacity utilization was decreasing ($c \downarrow$). We now turn to an analysis of these results.

Consider first regressions 1 to 3 of Table 2. These contain ω as a function of the rate of capacity utilization alone. The first overall impressions are that the R^2 are a little lower than in Table 1 and that the capacity variables do not perform particularly well. The coefficients of the capacity variables possess the correct signs but are significant at the 5 percent level only when the rate of capacity utilization is not subject to a subdivision into more than one series. There is no evidence of an asymmetrical response of the foreign domestic good ratio to variations of capacity utilization over the cycle (see equation (2)) nor is there any evidence that the relationship is non-linear (see equation (3)).

When the rate of capacity utilization is lagged it becomes increasingly insignificant and often exhibits the incorrect sign. For those variables which are also common to Table 1 the magnitude of their coefficients and the degree of significance of the estimates proved to be, in general, insensitive to the inclusion of the capacity utilization rate as a measure of the ω elements of the effective price. Consequently, we will not discuss these results in detail.

In regression equations 4, 5, and 6, the capacity utilization variable is combined with some of the variables suggested by the theory of Section II. In these circumstances the capacity utilization variable is usually insignificant although often marginally so. There is again no evidence of non-linearities or asymmetrical responses over the cycle.

These results suggest that little is to be gained from specifying ω as a function of the capacity utilization rate, or from introducing the rate of capacity utilization into the specification of the effective price. Furthermore, they add support to the view that it is not bottleneck and capacity constraints per se that affect the allocation of demand between foreign and

THE AMERICAN ECONOMIC REVIEW

TABLE 2—THE RATIO OF IMPORTS TO DOMESTIC GOODS PRODUCTION REGRESSED
AGAINST VARIOUS SPECIFICATIONS OF THE EFFECTIVE PRICE RATIO
Quarterly Data 1948–68 (*t*-values in parentheses)

Equations	1	2	3	4	5	6
Constant	-.3058 (.71)	-.2609 (.58)	-.4023 (.49)	.0823 (.20)	.1210 (2.84)	-.1385 (.33)
$\frac{P_1}{P_2}$.3751 (2.89)	.3660 (2.78)	.3801 (2.90)	.9398 (5.17)	.5348 (3.99)	.5117 (3.72)
<i>t</i>	.0010 (2.05)	.0009 (2.00)	.0011 (2.36)	.0007 (1.80)	.0010 (2.32)	.0010 (2.34)
$\frac{X_{2t-1}}{X_{1t-1}}$.7708 (10.85)	.7735 (10.74)	.7612 (10.69)	.6657 (10.12)	.6890 (9.55)	.7034 (9.49)
c_t	.1984 (2.00)			.0569 (.61)		.1989 (2.07)
$c \uparrow_t$.1867 (1.80)			.1486 (1.53)	
$c \downarrow_t$.1867 (1.78)			.1485 (1.51)	
c_{1t}			.2270 (1.21)			
c_{2t}			.2220 (1.21)			
c_{3t}			.2256 (1.25)			
l_{at}				.5490 (4.94)		
l_{at-1}				-.4196 (4.48)		
$\frac{P_{1t}}{P_{1t-1}}$					1.7551 (2.75)	1.7280 (2.45)
$\frac{P_{1t-1}}{P_{1t-2}}$					1.3250 (1.89)	
S.E.R.	.047	.046	.047	.041	.044	.046
R^2	.92	.92	.92	.94	.93	.92
<i>D</i>	1.50	1.51	1.20	1.24	.36	.74

Symbols: c_t = the rate of capacity utilization in the manufacturing sector; $c \uparrow_t$ = the rate of capacity utilization when the capacity utilization index is increasing; $c \downarrow_t$ = the rate of capacity utilization when the capacity utilization index is decreasing; c_{1t} = the rate of capacity utilization < 81 percent; c_{2t} = the rate of capacity utilization > 81 and < 89 percent; c_{3t} = the rate of capacity utilization > 89 percent.

Source: The capacity utilization indices are taken from the *Federal Reserve Bulletin*. *D* is a measure of serial correlation. All variables are expressed in logarithms, seasonally adjusted, and measured in constant prices.

domestic goods, but relative effective prices.

Preliminary Application of the Model to Disaggregated Data

Any attempt to apply the model to data at lower levels of aggregation runs into quite serious data problems. It is difficult to match each category of imports with the relevant category of domestic goods and, in general, data series of the price indices of imports, and for most of the proxy variables suggested in Section II, are unavailable. However, despite these problems it may be worthwhile to report very briefly on some preliminary results.²⁰

Table 3 lists the results of the first attempts to apply the model to lower levels of data aggregation. The categories were chosen to reduce the data problems, although they are considerable. Consider the first two categories, producer and consumer durables. The estimates of the magnitude and degree of significance of the coefficients of the relative price term, the lagged dependent variable, and the time trend are very similar to those of the aggregate data. The coefficient on the lagged dependent variable is still very significant and the response elasticity remains approximately .3. The coefficient on the time trend is again positive and significant. The short-run elasticity of substitution is significant and approximately .6. The proxy variables for the ω elements of the effective price possess the correct sign and are generally significant.

The least satisfactory results occurred in the third category, consumer nondurables. The price elasticity of substitution is insignificant, although it possesses the

correct sign when the waiting time and its rate of change are added to the equation. Both waiting time variables possess the correct sign, but only one is significant at conventional levels.

As a generalization it could be said that the results support the hypotheses that the allocation of resources between domestic and foreign goods is responsive to variations in the non-price rationing variables discussed earlier. However, the model outlined in Section II does not perform quite as well at this level of aggregation. Whether this stems wholly from the inadequacy of the data (see Table 3 for the data description) or from the inadequacy of the model is impossible to say at this moment. There is obviously a need for considerably more work on the disaggregated data series.

IV. Conclusions

1. This paper is primarily concerned with two hypotheses. First, in the short run producers do not adjust prices to meet fluctuations of demand, but use other rationing devices, such as waiting times, credit terms, and even the enthusiasm with which they react to the enquiries of prospective customers. Second, the allocation of demand between domestic and foreign goods is responsive to these short-run rationing devices. A test of these hypotheses against the behavior of the ratio of U.S. imports to the production of domestic goods from the 1948-68 data reveals that the evidence supporting these hypotheses is quite considerable.

2. As a result of data difficulties these short-run rationing variables cannot be measured directly. The model of the firm outlined in Section II suggests that the quoted waiting time might be measured by the estimated delivery lag and its rate of change. Other proxy variables include the inventory level and its rate of change, and the rate of change of prices. When

²⁰ The theory of the previous sections can be applied to the disaggregated data by the extension of the utility tree to further sub-levels. Where the commodity in question is a producer good the utility function is replaced by a production function and the same analysis is applicable.

TABLE 3—IMPORT DEMAND FUNCTIONS FOR VARIOUS CATEGORIES OF IMPORTS
Quarterly Data 1953-68 (*t*-values in brackets)

Equations	Producer Durables				Consumer Durables			Consumer Nondurables	
	1	2	3	4	1	2	3	1	2
Constant	-2.1928 (3.92)	-2.2307 (4.25)	-2.1353 (4.02)	-2.1800 (4.27)	-1.2368 (3.47)	-1.2180 (3.94)	-1.0239 (2.85)	-1.5795 (3.01)	-1.0561 (1.90)
$\frac{P_{1t}}{P_{2t}}$.6390 (2.87)	.7376 (3.49)	.6155 (2.88)	.6962 (3.35)	.6029 (2.42)	.6460 (2.30)	.5430 (2.23)	-.1083 (.20)	.3324 (.62)
<i>t</i>	.0087 (3.39)	.0088 (3.98)	.0077 (2.52)	.0082 (2.78)	.0056 (3.39)	.0058 (2.94)	.0045 (2.71)	.0077 (2.82)	.0107 (3.02)
$\frac{X_{2t-1}}{X_{1t-1}}$.6374 (7.02)	.6337 (7.44)	.6574 (7.03)	.6467 (7.20)	.6755 (7.36)	.6737 (7.29)	.7295 (7.91)	.7168 (7.63)	.6483 (6.28)
$\frac{P_{1t-1}}{P_{1t-2}}$		3.1733 (2.89)		2.5565 (2.32)			2.0455 (2.14)		
l_{ot-1}						.2583 (1.20)			.4413 (2.77)
l_{ot-2}			.2967 (2.51)	.2433 (2.10)		-.2526 (1.31)			-.2109 (1.32)
l_{ot-3}			-.3267 (2.77)	-.2590 (2.22)					
S.E.R.	.059	.056	.056	.054	.056	.056	.055	.074	.070
R^2	.98	.99	.98	.99	.97	.97	.97	.97	.98
<i>D</i>	1.06	.25	1.06	.64	.72	1.20	.48	.96	.96

Symbols are defined in Table 1: Dependent Variables: Producer durables: Imports of capital equipment divided by Producers' durable equipment; Consumer durables: Imports of consumer durables, manufactured, divided by consumer expenditure on durables minus imports of consumer durables and consumer expenditure on automobiles and parts. Consumer nondurables: Imports of consumer nondurables, manufactured, divided by consumer expenditure on nondurable goods minus imports of consumer durables and consumption expenditure on food and beverages. Price Indices: Implicit price deflators of imports, producer durable equipment, personal consumption expenditure (durable goods), personal consumption expenditure (nondurable goods). Delivery lags: Calculated from unfilled orders, durable goods industries, and unfilled orders, nondurable goods industries. All data are taken from the *Survey of Current Business*. Data are seasonally adjusted, in constant prices and expressed as logarithms.

these variables are included in the demand function each possesses the predicted sign and magnitude and is statistically significant. A simple measure of the magnitude of domestic capacity constraints—the rate of utilization of capacity in the manufacturing sector—does not perform particularly well.

3. The inclusion of these proxy variables to measure variations in effective prices leads to an estimate of the price elasticity

of substitution between foreign and domestic goods which is larger, and statistically more significant, than estimates derived from traditional specifications of demand equations. The estimates of the elasticity of substitution imply a short-run price elasticity for all imports of between .5 and unity and a long-run elasticity of demand of approximately three.

4. After allowance has been made for past changes in relative prices and the

non-price rationing variables, there has been a steady trend towards increasing the proportion of goods consumed which are imported.

5. Preliminary estimation of the model suggests that the above hypotheses may be supported when the data is further disaggregated but there remains considerable work to be done at lower levels of the data aggregation to derive conceptually adequate data and to test the model more fully.

6. In conclusion, the evidence presented here would suggest that on both theoretical and empirical levels more attention could be given to short-run disequilibrium situations and to variables which have not been traditionally included in demand and supply equations. Our model represents a step in this direction. In particular the theoretical model would benefit from a dynamic specification.

REFERENCES

- R. J. Ball, J. R. Eaton, and M. D. Steuer, "The Relation between United Kingdom Export Performance in Manufactures and the Internal Pressure of Demand," *Econ. J.*, Sept. 1966, 76, 501-19.
- R. J. Ball and K. Marwah, "The U.S. Demand for Imports, 1948-58," *Rev. Econ. Statist.*, Nov. 1962, 44, 395-401.
- W. H. Branson, "A Disaggregated Model of the U.S. Balance of Trade," *Staff Economic Papers*, Board of Governors of the Federal Reserve System.
- F. Brechling and J. N. Wolfe, "The End of Stop-Go," *Lloyds Bank. Rev.*, Jan. 1965, 73, 23-30.
- P. A. David and Th. Van De Klundert, "Biased Efficiency Growth and Capital Labour Substitution in the U.S., 1899-1960," *Amer. Econ. Rev.*, June 1965, 55, 355-94.
- J. Durbin, "Testing for Serial Correlation in Least Squares Regression when some of the Regressors are Lagged Dependent Variables," *Econometrica*, forthcoming.
- O. Eckstein and G. Fromm, "The Price Equation," *Amer. Econ. Rev.*, Dec. 1968, 58, 1159-83.
- J. Enzler, "Revised Indexes of Manufacturing Capacity and Capacity Utilization," *Fed. Res. Bull.*, July 1967, 53, 1096-98.
- W. Godley and J. Shepherd, "Forecasting Imports," *Nat. Inst. Econ. Rev.*, Aug. 1965, 33, 35-42.
- S. M. Goldman and H. Uzawa, "A Note on Separability in Demand Analysis," *Econometrica*, July 1964, 32, 387-98.
- Z. Griliches, "Distributed Lags: A Survey," *Econometrica*, Jan. 1967, 35, 16-49.
- K. Lancaster, "A New Approach to Consumer Theory," *J. Polit. Econ.*, Apr. 1966, 74, 132-57.
- W. Lederer, E. M. Parrish, and S. Pizer, "The Balance of International Payments: Fourth Quarter and Year 1965," *Surv. Curr. Bus.*, U.S. Department of Commerce, Mar. 1966, 46, 16-28.
- E. Malinvaud, *Statistical Methods of Econometrics*, Amsterdam 1966.
- E. S. Phelps, "The New Microeconomics in Inflation and Employment Theory," *Amer. Econ. Rev. Proc.*, May 1969, 59, 147-60.
- R. Rhomberg and L. Boissonneault, "The Foreign Sector" in J. S. Duesenberry, ed., *Brookings Quarterly Econometric Model of the United States*; Chicago 1965, 375-408.
- P. A. Samuelson, *Foundations of Economic Analysis*, Cambridge, Mass. 1947.
- K. Sato, "A Two Level Constant Elasticity of Substitution Production Function," *Rev. Econ. Stud.*, Apr. 1967, 34, 201-18.
- M. D. Steuer, R. J. Ball, and J. R. Eaton, "The Effect of Waiting Times on Foreign Orders for Machine Tools," *Economica*, Nov. 1966, 33, 387-403.
- L. C. Thurow, "A Fiscal Policy Model of the United States," *Surv. Curr. Bus.*, U.S. Department of Commerce, June 1969, 49, 45-64.
- Surv. Curr. Bus.*, U.S. Department of Commerce, various issues.
- U.S. Congress Joint Economic Committee Subcommittee on Economic Statistics, *Hearings*, Jan 24, 1961, Washington 1961.

Cropsharing Tenancy in Agriculture: A Theoretical and Empirical Analysis

By P. K. BARDHAN AND T. N. SRINIVASAN*

Cropsharing tenancy is one of the earliest forms of production organization in agriculture. It is still a matter of considerable importance in peasant agriculture in many countries. While since Adam Smith much of the discussion¹ in the literature has been on the comparative efficiency of this with other forms of land institutions, it is less easy to find an analysis of the economic factors that may explain variations in the incidence of share-cropping in different areas. Needless to say, land institutions are shaped by diverse historical, political, and sociological factors peculiar to different regions; but an economist persists in believing that the "relations of production" ultimately reflect some of the basic economic "forces of production." The object of this paper is first to identify some of the relevant economic factors in terms of a simple theoretical framework and then check for their actual significance in an empirical analysis of interregional cross-section data pertaining to Indian agriculture.

In Section I, we discuss the basic theoretical model and the properties of equilibrium; in Section II we analyze the comparative static results with respect to variations in the wage rate, and in Section III, those with respect to parameters repre-

senting different kinds of technical progress; in Section IV we report the empirical results confirming some of the hypotheses in the theoretical sections; in Section V we comment on the problem of cost-sharing by landlords, and in the concluding Section VI, we briefly touch on the problem of uncertainty.

I. The Model and the Properties of Equilibrium

Let us take agricultural output to be a single homogeneous commodity whose production requires, for the time being, only land and labor. A landless person has the option to lease in some land and cultivate it with his labor or to work as wage-laborer on somebody else's farm or in non-agricultural occupations. In the former case he gets the output from the leased-in land after payment of a percentage share to the landlord and in the latter case, he receives a wage rate per unit of labor. We shall assume this wage rate to be given. The level of the agricultural wage rate depends on a whole host of factors outside the agricultural sector; for example, given other things, one expects the rural wage rate to be higher in West Bengal than in, say, Orissa, simply because by any index the former State is much more industrialized and urbanized. In a simplified model it seems better to assume that the wage is given as a parameter to the agricultural sector than to assume that it is completely determined by endogenous factors within the sector. So if our landless person consumes all of his income, his consumption is given by

* The authors are professors of economics at the Indian Statistical Institute, New Delhi. We have benefitted from suggestions for improvement by Amartya Sen and the editor and from the discussion in seminars at M.I.T., Delhi School of Economics and Vishwabharati at Santiniketan, where an earlier version of the paper was presented. All errors are, of course, ours alone.

¹ See a brief summary of it in D. Gale Johnson.

$$(1) \quad C^1 = (1 - r)F(H, l_1) + l_2w,$$

where r is the proportionate share of output paid as rent to the landlord, F is the production function, H is the amount of land hired by the tenant, l_1 is the amount of labor he devotes to the tenant farm, l_2 is the amount of labor he devotes to wage-paid occupations, and w is the given wage rate. We assume the production function is strictly concave and that its cross-partial derivatives are positive (i.e., the marginal product of one factor is an increasing function of the use of the other factor).

If we assume, without loss of generality, that the sharecropper has only one unit of labor available to him, then he devotes l_1 amount of time to his tenanted land, l_2 to outside work, and $(1-l_1-l_2)$ amount to leisure. If he is a maximizer, he maximizes a utility function

$$U^1(C^1, 1 - l_1 - l_2),$$

subject to the budget constraint in equation (1). We assume that the utility function is strictly concave and that neither consumption nor leisure is an inferior good.

If we confine ourselves to an interior maximum,² the necessary conditions are:

$$(2) \quad (1 - r)F_1 = 0$$

$$(3) \quad (1 - r)F_2 - w = 0$$

$$(4) \quad U_1^1 w - U_2^1 = 0$$

As usual, subscript 1 refers to the first derivative with respect to the first argument and subscript 2 to the second argument.

² Empirically this is the more interesting case. There is plenty of evidence in India that tenants are also part-time agricultural laborers or that primarily agricultural labor households have income from tenant cultivation. On the landlord side also, simultaneous leasing-out of land and self-cultivation of the rest is quite frequent.

We might briefly comment on these conditions. Equation (2) implies that the sharecropper will tend to lease in more and more land until the marginal product of land is driven to zero.³ Equation (3) refers to the usual result that a sharecropper will not use as much of a variable factor (in this case, labor) as an owner-farmer would; this leads to the standard case of inefficiency in sharecropping that economists from Adam Smith through Marshall have noted. Equation (4) also is a standard result: the optimum allocation of time between leisure and earning income implies that the ratio of their marginal utilities should be equalized to the wage rate.

The Jacobian matrix⁴ $[a_{ij}]$ of equations (2), (3), and (4) will have the following elements:

$$a_{11} = (1 - r)F_{11}; \quad a_{12} = (1 - r)F_{12};$$

$$a_{13} = a_{23} = a_{31} = 0; \quad a_{21} = (1 - r)F_{21};$$

$$a_{22} = (1 - r)F_{22};$$

$$a_{32} = a_{33} = \frac{U_2^1}{U_1^1} \left[U_{11}^1 \frac{U_2^1}{U_1^1} + U_{22}^1 \frac{U_1^1}{U_2^1} - 2U_{21}^1 \right]$$

Totally differentiating (2), (3), and (4) with respect to r (with w given), we get

$$(5) \quad [a_{ij}] \begin{bmatrix} \frac{dH}{dr} \\ \frac{dl_1}{dr} \\ \frac{dl_2}{dr} \end{bmatrix} = \begin{bmatrix} 0 \\ F_2 \\ F \left(U_{11}^1 \frac{U_2^1}{U_1^1} - U_{21}^1 \right) \end{bmatrix}$$

³ Particular attention to this condition was drawn by Johnson.

⁴ The Jacobian $[a_{ij}]$ is negative, since with strictly concave production function, $F_{11}F_{22} > (F_{12})^2$, and with no inferior goods in the utility function, it is easy to show that $a_{33} < 0$.

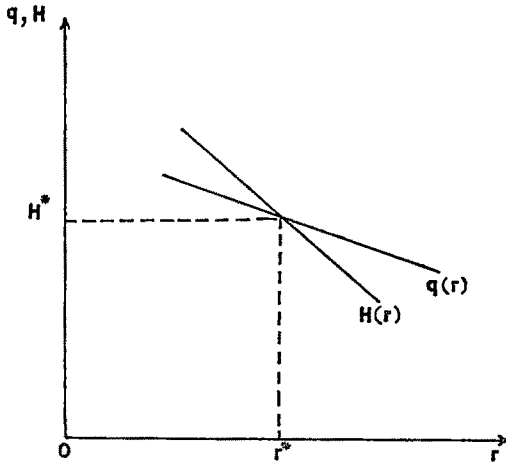


FIGURE 1. DEMAND AND SUPPLY IN THE MARKET FOR LAND LEASES

Using Cramer's Rule for solving simultaneous equations we get:

$$(6) \quad \frac{dH}{dr} = \frac{-F_2 F_{12}}{(1-r)[F_{11} F_{22} - (F_{12})^2]} < 0$$

$$(7) \quad \frac{dl_1}{dr} = \frac{F_2 F_{11}}{(1-r)[F_{11} F_{22} - (F_{12})^2]} < 0$$

Equations (6) and (7) imply that under our assumptions the amount of land leased in by the tenant as well as the amount of labor devoted to it, is a declining function of the rental share. Figure 1 depicts $H(r)$, the demand function for land to be leased in by sharecroppers.

Let us now look at the supply side. The land owner has the option to cultivate land with the use of own and hired labor or to lease out land to sharecroppers. In the former case, he has to pay hired labor at the given wage rate and in the latter, he gets only a share in the tenant's output. Assuming again that all of income is consumed, his consumption is

$$(8) \quad C^2 = G(1-q, x+y) - wx + rF(q, L)$$

where G is the production function in his own cultivation; assuming that the land

owner has one unit of land, q is the proportion (or amount) of land he leases out to sharecroppers retaining $(1-q)$ for self-cultivation; x is the amount of hired labor and y is the amount of own labor (for simplicity we are assuming that these two types of labor are qualitatively similar), and L is the amount of labor the sharecropper puts in on the leased-out land.

If, again without loss of generality, we assume that the land owner has one unit of time available to him, then $(1-y)$ is the amount of time he takes out as leisure. He thus maximizes a utility function

$$U^2(C^2, 1-y)$$

subject to the budget constraint in equation (8). The properties of G and U^2 functions are assumed to be similar to those of F and U^1 functions we have encountered before. In the maximization process we should note that the land owner can vary q , x , and y , but he cannot choose L ; it is up to the sharecropper to decide how much labor he would like to use on the tenanted land.

Once again confining ourselves to an interior maximum the necessary conditions are:

$$(9) \quad G_1 - rF_1 = 0$$

$$(10) \quad G_2 - w = 0$$

$$(11) \quad U_1^2 G_2 - U_2^2 = 0$$

Equation (9) implies that the marginal product of land in self-cultivation should be equal to the share of the marginal product of land that land owner receives from leased-out land. Equation (10) implies that the marginal product of labor in self-cultivation should be equal to the wage rate. Equation (11) implies that the optimum allocation of time between leisure and earning income implies an equality of the ratio of their marginal utilities to the marginal productivity of labor.

The Jacobian matrix⁵ $[b_{ij}]$ of equations (9), (10), and (11) will have the following elements:

$$b_{11} = G_{11} + rF_{11}; \quad b_{12} = b_{13} = G_{12};$$

$$b_{21} = G_{21}; \quad b_{22} = b_{23} = G_{22};$$

$$b_{31} = U_1^2 \cdot G_{21}; \quad b_{32} = U_1^2 \cdot G_{22};$$

$$b_{33} = U_1^2 \cdot G_{22} + \frac{U_2^2}{U_1^2} \left[U_{11}^2 \frac{U_2^2}{U_1^2} + U_{22}^2 \frac{U_1^2}{U_2^2} - 2U_{21}^2 \right]$$

Totally differentiating (9), (10), and (11) with respect to r (with w and L given), we get

$$(12) \quad [b_{ij}] \begin{bmatrix} \frac{dq}{dr} \\ \frac{dx}{dr} \\ \frac{dy}{dr} \end{bmatrix} = \begin{bmatrix} F_1 \\ 0 \\ F \left(U_{11}^2 \frac{U_2^2}{U_1^2} - U_{21}^2 \right) \end{bmatrix}$$

Using Cramer's Rule,

$$(13) \quad \frac{dq}{dr} \Big|_{L=L_0} = \frac{-F_1 \cdot G_{22}}{[G_{11}G_{22} - (G_{12})^2 + rF_{11} \cdot G_{22}]}$$

Similarly,

$$(14) \quad \frac{dL}{dr} \Big|_{r=r_0} = \frac{-rF_{12} \cdot G_{22}}{[G_{11} \cdot G_{22} - (G_{12})^2 + rF_{11} \cdot G_{22}]} > 0$$

Taking the demand and supply side (for land-leases)⁶ together, we can now write

⁵ Again the Jacobian in this case is negative with strictly concave production function and no inferior goods in the utility function.

⁶ In this paper we ignore the market for land buying and selling and concentrate our attention on land renting.

the equilibrium conditions⁷ as

$$(15) \quad H(r) = q(r, L)$$

$$(16) \quad L = l_1(r)$$

Equations (15) and (16) may be rewritten as one condition:

$$(17) \quad H(r) = q(r, l_1(r))$$

From (6) we know that $dH/dr < 0$ and from (7), (13), and (14) we can now write that around equilibrium

$$(18) \quad \frac{dq}{dr} = \frac{dq}{dr} \Big|_{L=L_0} + \frac{dq}{dL} \Big|_{r=r_0} \left(\frac{dl_1}{dr} \right) = \frac{dq}{dL} \Big|_{r=r_0} \left(\frac{dl_1}{dr} \right) < 0,$$

since the equilibrium $F_1 = 0$.

So around equilibrium the supply of land leased out is also a *declining* function of the rental share. But it can be easily checked from (6), (7), and (14) that around equilibrium

$$(19) \quad \frac{dH}{dr} - \frac{dq}{dr} < 0$$

so that in Figure 1, the $H(r)$ curve must intersect the $q(r)$ curve from above. Using the standard tools of Walrasian stability analysis the equilibrium is unique and stable. H^* is the equilibrium amount (or percentage) of land leased out to sharecroppers and r^* is the competitively determined rental share.⁸

⁷ Note that in equilibrium x is not equal to l_2 ; otherwise the agricultural wage rate would have been endogenously determined.

⁸ We have so far ignored any form of land lease at a given rental rate per unit of land (rather than at a given rental share of output). Johnson has suggested that where there is this form of renting simultaneously with sharecropping the equilibrium conditions are different from those under sharecropping alone. We shall, however, show here that both these forms of land-lease cannot coexist in our model.

(over)

Before we pass on to our comparative-static analysis, let us briefly comment on some points raised in a recent paper on sharecropping by S.N.S. Cheung (1968). His paper, like the present one, tries to change the analysis of sharecropping from its usual partial-equilibrium framework to a simple general equilibrium footing, although the number of choices available to the economic agents are larger in the present paper (for example, Cheung does not give landlords the option to self-cultivate). Some of the implications of our model seem to be in contradiction with those of Cheung. One of his major conclusions is that the prevailing impression of inefficient allocation of resources under sharecropping is wrong. This does not seem to tally with our equation (3) where, clearly, the sharecropper stops short of equalizing the marginal product of labor to the wage rate.

Suppose (1) is rewritten as

$$C^1 = (1 - r)F^1(H_1, l_1) + F^2(H - H_1, l_2) - R(H - H_1) + w(1 - l_1 - l_2),$$

where H_1 and l_1 are the amounts of land and labor, respectively, under sharecropping, F^1 is the production function on the sharecropped farm, F^2 is the production function on the land leased under the other form of renting, R is the given rental rate per unit of such land; and for simplification, let us ignore leisure. One of the necessary conditions for maximizing C^1 is

$$rF_1^1 \leq 0 \quad \text{as } H_1 = \begin{cases} 0 \\ \epsilon \\ H \end{cases} (0, H)$$

Similarly, landlord's consumption is

$$C^2 = G(1 - q, 1 + x) - wx + rF^1(q_1, L_1) + R(q - q_1)$$

where q_1 is the amount of land leased out under sharecropping and once again we ignore leisure. One of the necessary conditions of maximizing C^2 is

$$rF_1^1 - R \geq 0 \quad \text{as } q_1 = \begin{cases} 0 \\ \epsilon \\ q \end{cases} (0, q)$$

Now putting these two necessary conditions together we can now say that $H_1 = q_1 \in (0, q)$ leads to a contradiction, for then $F_1^1 = 0$ which means $R = 0$ which means no land will be rented out under the alternative form of renting.

So it seems in our model all of leased land is either under sharecropping or under a given rental rate per unit of land. We have chosen in this paper to devote our analysis only to the former problem.

The difference lies in the kind of maximization process Cheung carried out in Section III, pp. 1113-14 of his paper. There he maximizes only from the landlord's point of view,⁹ whereas in this paper we determine the demand side from maximization by the tenant, just as the supply side is determined from landlord's maximizing decision. Further, Cheung's landlords can freely choose the amount of labor the sharecropper would devote to the tenant farm whereas in our paper this is decided by the sharecropper alone. It seems to us that unless the share contract explicitly contains stipulations regarding the intensity of use of labor by the sharecropper, the better assumption is to leave the decision to the sharecropper.

It is also difficult to understand why Cheung's landlord maximizes with respect to r , the rental share. Cheung seems to think that in the traditional analysis of sharecropping, "the writers fail to realize that the percentage shares and area rented under share tenancy are not mysteriously fixed but are competitively determined in the market"; yet in his analysis the landlord himself decides about the level of r . In a competitive situation one would have thought that each atomistic landlord and each atomistic tenant would take r as given, and out of their aggregate maximizing decisions r is determined in the market, as in our paper. It seems that only when the landlord has monopoly power in the market for land-leases can he choose the level of r .

Cheung also introduces an extra decision variable, viz. the number of parcels into which the landlord can subdivide his land leasing out each to a different tenant. In

⁹ The only way the tenant's decision enters in any way in Cheung's model is that the tenant's income from cultivation should be equal to the amount he would have earned had he been a wage laborer, i.e., in our notation, $wl_1 = (1-r)F$. We have seen above that if the tenant can devote part of his labor to tenant-cultivation and part to wage-employment, this changes to $w = (1-r)F_1$.

that case there is no reason why he should ignore a similar decision variable on the part of the tenant, i.e., the tenant can also choose the number of landlords from each of whom he leases in a parcel of land.

In our model above we have ignored these two decision variables. Let us briefly mention here how one may incorporate them in our model. If n is the number of landlords from whom the tenant leases in a parcel of land, equation (1) has to be rewritten as:

$$(20) \quad C^1 = n(1 - r)F(h, l_1) + l_2w,$$

where h is the size of each tenanted parcel and l_1 is the amount of labor devoted to it. The utility function has to be rewritten as

$$U^1(C^1, 1 - l_1n - l_2)$$

When this function is maximized with respect to h , l_1 , and l_2 , we get exactly the same necessary conditions as (2), (3), and (4). It is to be noted particularly that the inefficiency in labor use under sharecropping that was implied by (3) remains valid even in this case.

Differentiating U^1 with respect to n , we get

$$(21) \quad \frac{\partial U^1}{\partial n} = (1 - r)U_1^1[F - F_2 \cdot l_1] > 0$$

This implies that the tenant continuously gains from parcellization. This is expected because of the diminishing returns to scale implied by our strictly-concave production function. For any meaningful analysis one should impose an upper bound on the number of feasible parcels, say \bar{n} , and if we assume that in our model n will always be equal to \bar{n} .

Similarly, from the landlord side also it is easy to show that a strictly-concave production function implies

$$\frac{\partial U^2}{\partial m} > 0,$$

when m is the number of parcels into which the landlord can subdivide his land leasing out each to a different tenant. Once again we shall assume an upper bound, \bar{m} , which will be reached.

In the rest of the analysis we shall, therefore, again ignore the problem of parcellization.

II. Parametric Shifts in the Wage Rate

In the model above we have taken the wage rate as exogenously given to the agricultural sector. Let us now find out the implications of parametric shifts in the wage rate.

Differentiating with respect to w in equations (2), (3), and (4), keeping r fixed for the moment, we get

$$(22) \quad [a_{ij}] \begin{bmatrix} H_w \\ l_{1w} \\ l_{2w} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -U_1^1 - l_2 \left(U_{11}^1 \frac{U_2^1}{U_1} - U_{21}^1 \right) \end{bmatrix}$$

where

$$H_w = \frac{\partial H}{\partial w} \Big|_{r=r_0}, \quad l_{1w} = \frac{dl_1}{dw} \Big|_{r=r_0}, \text{ etc.}$$

Using Cramer's Rule again,

$$(23) \quad H_w = \frac{-F_{12}}{(1 - r)[F_{11} \cdot F_{22} - (F_{12})^2]} < 0$$

$$(24) \quad l_{1w} = \frac{F_{11}}{(1 - r)[F_{11} \cdot F_{22} - (F_{12})^2]} < 0$$

Equations (23) and (24) imply that for a given rental share, the higher the wage rate the lower is the tenant's demand for leasing-in land (since wage employment is more attractive now) and also the amount of labor devoted to tenanted land.

Differentiating with respect to w in

equations (9), (10), and (11), keeping r and L fixed for the moment, we get

$$(25) \quad [b_{ij}] \begin{bmatrix} q_w \\ x_w \\ y_w \end{bmatrix} = \begin{bmatrix} 0 \\ x \left(U_{11}^2 \frac{U_2^2}{U_1^2} - U_{21}^2 \right) \end{bmatrix}$$

We may solve for

$$(26) \quad q_w = \frac{G_{12}}{[G_{11} \cdot G_{22} - (G_{12})^2 + r F_{11} \cdot G_{22}]} > 0$$

This means that given r and L , a rise in the wage rate for hired labor induces land-owners to lease out more land.

But in equilibrium with a rise in the wage rate r cannot remain fixed. Let us rewrite equilibrium condition (17) as

$$(27) \quad H(r, w) = q(r, w, l_1(r, w))$$

Using (14), (19), (23), (24), and (26) we can get equation (28), which means that a parametric rise in the wage rate brings down the equilibrium rental share.

Let us now find out how such a parametric shift in the wage rate affects the equilibrium amount (or percentage) of land leased out under cropsharing.

From (27), and using (6), (23), and (28), we obtain (29) which implies that a

parametric rise in the wage rate leads to a *rise* in the equilibrium amount (and percentage) of land leased out under crop-sharing. In our empirical analysis of cross-sectional data, we shall accordingly try to test the hypothesis that a region with a higher wage rate will tend to have a larger incidence of cropsharing tenancy.

III. Technical Progress

In all of our analysis above we assumed that output depends on the use of labor and land and our production functions did not admit of any technical change. We shall now introduce parameters representing different kinds of technical progress.

Let us rewrite tenant's consumption equation (1) as

$$(30) \quad C^1 = (1 - r)\rho F(\mu H, \lambda l_1) + l_2 w$$

where ρ , μ , and λ are multiplicative technical progress parameters. We shall take three types of technical progress corresponding to types familiar in the growth literature:

- a) Land-augmenting technical progress, when ρ and λ are constant but μ shifts;
- b) Hicks-Neutral technical progress, when μ and λ are constant but ρ shifts and
- c) Labor-augmenting technical pro-

$$(28) \quad \frac{dr}{dw} = \left\{ (1 - r)[F_{11} \cdot F_{22} - (F_{12})^2][G_{11} \cdot G_{22} - (G_{12})^2 + r F_{11} \cdot G_{22}] \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} \\ \cdot [F_{12}\{G_{11} \cdot G_{22} - (G_{12})^2\} + (1 - r)G_{12}\{F_{11} \cdot F_{22} - (F_{12})^2\}] < 0$$

$$(29) \quad \frac{dH}{dw} = H_w + \frac{\partial H}{\partial r} \cdot \frac{dr}{dw} \\ = - \left\{ (1 - r)[F_{11} \cdot F_{22} - (F_{12})^2][G_{11} \cdot G_{22} - (G_{12})^2 + r F_{11} \cdot G_{22}] \right. \\ \cdot \left. \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} F_2 \cdot F_{12} \cdot G_{12} > 0$$

gress, when ρ and μ are constant but λ shifts. We shall assume that the same parameters obtain with respect to G , the landlord's self-cultivation function.

Taking all these parameters into account, the necessary conditions of tenant's utility maximization given by equations (2), (3), and (4) will have to be rewritten as

$$(31) \quad (1-r)\rho\mu \cdot F_1 = 0$$

$$(32) \quad (1-r)\rho\lambda F_2 - w = 0$$

$$(33) \quad U_1^1 w - U_2^1 = 0$$

Similarly, the landlord's utility maximization will involve

$$(34) \quad \mu\rho(G_1 - rF_1) = 0$$

$$(35) \quad \lambda\rho G_2 - w = 0$$

$$(36) \quad \lambda\rho U_1^2 G_2 - U_2^2 = 0$$

We shall now use these equations for our comparative-static analysis.

Let us first take case a), i.e., land-augmenting technical progress, and assume for simplicity that $\rho=\lambda=1$. Using differentiation and Cramer's Rule as before it is easy to show from equations (31), (32), and (33) that

$$(37) \quad H_\mu = \frac{dH}{d\mu} \Big|_{r=r_0} = -\frac{H}{\mu} < 0$$

and

$$(38) \quad l_{1\mu} = \frac{dl_1}{d\mu} \Big|_{r=r_0} = 0$$

Similarly, from equations (34), (35), and (36),

$$(39) \quad q_\mu = \frac{dq}{d\mu} \Big|_{r=r_0} = \frac{\{\mu[G_{11}G_{22} - (G_{12})^2 + rF_{11}G_{22}]\}^{-1} \cdot [(1-q)\{G_{11}G_{22} - (G_{12})^2\} - qrF_{11} \cdot G_{22}]}{\mu[G_{11}G_{22} - (G_{12})^2 + rF_{11}G_{22}]\}^{-1} > 0$$

$$(40) \quad \frac{\partial q}{\partial L} = -rF_{12} \cdot G_{22}$$

The equilibrium condition (17) is now rewritten as

$$(41) \quad H(r, \mu) = q(r, \mu, l_1(r))$$

Using (19), (37), (39), and (41) we can get

$$(42) \quad \frac{dr}{d\mu} = \left\{ \mu[G_{11}G_{22} - (G_{12})^2 + rF_{11}G_{22}] \cdot \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} \cdot [G_{11}G_{22} - (G_{12})^2] < 0$$

In other words, the rental share is a declining function of the land-augmenting technical progress parameter.

Using (6), (7), (19), (37), (39), (40), and (42), we obtain equation (43) which shows that with land-augmenting technical progress the equilibrium amount (and percentage) of land leased out under crop-sharing goes up.

Let us now take case b), i.e., neutral technical progress and assume for simplicity that $\mu=\lambda=1$. From (31), (32), and (33),

$$(43) \quad \frac{dH}{d\mu} = - \left\{ \mu^2(1-r)[F_{11}F_{22} - (F_{12})^2][G_{11}G_{22} - (G_{12})^2 + rF_{11}G_{22}] \cdot \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} [G_{11}G_{22} - (G_{12})^2](1-q) > 0$$

$$(44) \quad H_\rho = \frac{dH}{d\rho} \Big|_{r=r_0} = \frac{F_2 F_{12}}{\rho [F_{11} F_{22} - (F_{12})^2]} > 0$$

$$(45) \quad l_{1\rho} = \frac{dl_1}{d\rho} \Big|_{r=r_0} = - \frac{F_2 \cdot F_{11}}{\rho [F_{11} F_{22} - (F_{12})^2]} > 0$$

Similarly, from (34) and (35), and (36),

$$(46) \quad q_\rho = \frac{dq}{d\rho} \Big|_{r=r_0} = \frac{-G_2 \cdot G_{12}}{\rho [G_{11} G_{22} - (G_{12})^2 + r F_{11} G_{22}]} < 0$$

The equilibrium condition (17) is now re-written as

$$(47) \quad H(r, \rho) = q(r, \rho, l_1(r, \rho))$$

From (14), (19), (44), (45), (46), and (47), we obtain equation (48) which shows that with a rise in the neutral technical progress parameter the rental share also goes up.

Using (6), (7), (19), (44), (45), and (46), we obtain equation (49) which shows that with neutral technical progress the equilibrium amount (and percentage) of land leased out under cropsharing goes down.

Let us finally take case c), i.e., labor-augmenting technical progress and assume for simplicity that $\rho = \mu = 1$. From (31), (32), and (33)

$$(50) \quad H_\lambda = \frac{dH}{d\lambda} \Big|_{r=r_0} = \frac{F_2 F_{12}}{\lambda [F_{11} \cdot F_{22} - (F_{12})^2]} > 0$$

$$(51) \quad l_{1\lambda} = \frac{dl_1}{d\lambda} \Big|_{r=r_0} = \frac{-[\lambda l_1 \{F_{11} F_{22} - (F_{12})^2\} + F_2 F_{11}]}{\lambda^2 [F_{11} F_{22} - (F_{12})^2]}$$

Similarly, from (34), (35), and (36),

$$(52) \quad q_\lambda = \frac{dq}{d\lambda} \Big|_{r=r_0} = \frac{-[G_2 G_{12} + r \lambda L F_{12} \cdot G_{22}]}{\lambda [G_{11} G_{22} - (G_{12})^2 + r F_{11} G_{22}]}$$

$$(53) \quad \frac{\partial q}{\partial L} = \frac{-r \lambda F_{12} G_{22}}{[G_{11} G_{22} - (G_{12})^2 + r F_{11} G_{22}]} > 0$$

The equilibrium condition (17) is now re-written as

$$(54) \quad H(r, \lambda) = q(r, \lambda, l_1(r, \lambda))$$

Using (19), (50), (51), (52), (53), and (54), one can show that $dr/d\lambda$ is exactly equal to the expression on the right-hand side of

$$(48) \quad \frac{dr}{d\rho} = - \left\{ \rho [F_{11} F_{22} - (F_{12})^2] [G_{11} G_{22} - (G_{12})^2 + r F_{11} G_{22}] \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} \cdot [G_2 G_{12} \{F_{11} F_{22} - (F_{12})^2\} + F_2 F_{12} \{G_{11} G_{22} - (G_{12})^2\}] > 0$$

$$(49) \quad \frac{dH}{d\rho} = \left\{ \rho [F_{11} \cdot F_{22} - (F_{12})^2] (1 - r) [G_{11} G_{22} - (G_{12})^2 + r F_{11} G_{22}] \cdot \left[\frac{\partial H}{\partial r} - \frac{\partial q}{\partial r} \right] \right\}^{-1} F_2 \cdot F_{12} G_2 \cdot G_{12} < 0$$

(48) with λ substituted for ρ . So we can say that with rise in the labor-augmenting technical progress parameter the rental share goes up.

Using (6), (7), (19), (50), (51), (52), (53), and (54), we can also show that $dH/d\lambda$ is exactly equal to the expression on the right-hand side of (49) with λ substituted for ρ . So we can say that with labor-augmenting technical progress the equilibrium amount (and percentage) of land leased out under cropsharing goes *down*.

We might mention here two examples of how the results in this section may be useful towards explaining observed phenomena. The first is in terms of U.S. agriculture and the second in terms of Indian agriculture. It is well known that sharecropping was quite prevalent in the Mississippi Delta at least up to the 1930's. It is sometimes maintained¹⁰ that since then large-scale mechanization substituting for hand-picking of corn and cotton has contributed to a rapid decline in sharecropping. If the introduction of labor-saving technical equipments is regarded as a case (instead of factor substitution along the same isoquant) of labor-augmenting technical progress (i.e., of an increase in the effective supply of labor), then this phenomenon in the southern United States agriculture is consistent with our conclusion above that labor-augmenting technical progress leads to a fall in the incidence of sharecropping.

In Indian agriculture, irrigation may largely be regarded as a factor bringing about land-augmenting technical progress. By improving land-productivity and facilitating double cropping, irrigation serves to increase the effective supply of land: in efficiency units a piece of irrigated land may be regarded as a multiple of a piece of unirrigated land of same

acreage. In terms of the conclusion of our crude model regarding land-augmenting technical progress we then expect that with an increase in the importance of irrigation, given other things, the incidence of sharecropping should *increase*. In terms of a cross-sectional study, this means that we expect a larger percentage of area to be under sharecropping in better irrigated regions. In our empirical analysis with Indian data we have tried to test this hypothesis.

IV. Empirical Analysis

On the basis of our comparative-static analysis in Sections II and III, we expect a positive correlation between the importance of cropsharing tenancy in India and the agricultural wage rate or irrigation. In this section, we try to test it with the help of cross-sectional data across States and across some villages in India. Our dependent variable is the percentage of operated area leased in under cropsharing tenancy (we shall call it S), which represents an index of the incidence of sharecropping. Our main independent variables are w , the agricultural wage rate, and I , the percentage of operated area irrigated. Needless to say, in the real world S will depend on various other economic and particularly noneconomic factors. But our purpose is only to test if we have been able to identify at least two major economic factors explaining variations in the incidence of sharecropping and if the signs of the regression coefficients for these two variables are as expected from our comparative-static analysis. Since we do not claim to have identified all the major explanatory variables, we are not unduly perturbed if our R^2 in the subsequent analysis is not always very high. Besides, tenancy data, particularly in situations where there is a lot of land reform legislation at least in the statute books, may be

¹⁰ See, for example, Richard Day.

quite treacherous sometimes, and even extremely good statistical fits to such data may be suspect. Regarding the tenancy figures we shall be using we may indicate here our general belief that they are more likely to be underestimates of the actual incidence of tenancy than otherwise; we proceed on the inevitable assumption that this under-reporting bias is fairly uniformly distributed across regions.

Let us first report the results on the State-level data for the year 1960-61. On the basis of National Sample Survey data for thirteen States, we had the following least-squares estimate for a log-linear equation:

$$(55) \log S = 2.967 + 1.428 \log w^{**}$$

(9.9) (0.64)

$$+ 0.574 \log I^{**} - 2.663 \log B; \quad R^{2**} = .602$$

(0.23) (2.4)

Here and later we use ** for significant at 5 percent level and * for significant at 1 percent level.

In (55) the regression coefficients for w and I are significant (at 5 percent level) and of expected positive signs. Let us explain the variable B . In all of our analysis above we assumed that the market for land-lease is competitive. But the Indian land market is very much imperfect. In equation (55) we have taken B as a very crude index for the bargaining power of land-owners: It is the Lorenz concentration index of land ownership in the rural areas of each State. The assumption is that the larger is this concentration index the stronger is the bargaining power of the landlords in the land market.

What sign do we expect for the regression coefficient of B ? Without going through detailed analysis we may just invoke a familiar result of the theory of monopoly. A monopolist tends to restrict sales of output below the competitive level. Similarly, one may expect that a landlord having a strong bargaining power in the land market

will restrict leasing out below the level that a competitive market will reach. On this ground we expect, other things remaining the same, a negative correlation between S and B . In (55) the coefficient for B is not very significant (it is significant only at 30 percent level), but the sign is negative.

The multiple correlation coefficient, R , for the estimate as a whole is significant at 5 percent level.

Next we report results on the basis of forty villages covered by farm management surveys of the Government of India, Ministry of Food and Agriculture: ten in Ferozepur (Punjab), ten in Amritsar (Punjab), ten in Sambalpur (Orissa), and ten in West Godavari (Andhra Pradesh). The least-square estimate in this case is

$$(56) \log S = -0.603 + 1.452 \log w^{**}$$

(0.914) (0.710)

$$+ 0.691 \log I^{*}; \quad R^{2*} = .35$$

(0.254)

Once again coefficients for w and I are quite significant and of expected positive signs. For lack of data we could not use the variable B in this case. Although R^2 is not very high for the estimate, the multiple correlation coefficient is significant at 1 percent level.

We have also used, though somewhat less successfully, the data for thirty-nine villages in Punjab and Western U.P. covered by surveys of the Delhi University Agro-Economic Research Centre. The least-square estimate that gives the best fit is

$$(57) S = 6.609 + 9.804 w^{*} + 0.026 I$$

(18.0) (2.5) (0.066)

$$- 9.369 B; \quad R^{2*} = .386$$

(18.0)

The coefficients for w , I , and B (concentration index of land ownership) are all of expected signs. But it is significant only for w . The multiple correlation coefficient is significant at 1 percent level.

V. Cost-Sharing by Landlords

In this section we comment on the problem of cost-sharing by landlords in the context of cropsharing. Let us suppose that apart from land and labor there is a third input to production, fertilizers. The landlord may or may not share in the cost incurred by the tenant on fertilizers. We shall discuss the equilibrium relationship between crop share and cost share.

Let the tenant's production function be $F(H, N, l)$ where H and l denote land and labor inputs and N the input of fertilizer. We shall assume that F is strictly concave and exhibits diminishing returns to scale. Let p be the price per unit of N in terms of output. Let r denote the landlord's share in output and β his share in fertilizer cost. Let w be the externally given wage rate that the tenant can earn as a wage laborer. Let the total labor that the tenant can render be unity (for simplification, we ignore leisure). Then the tenant's income (consumption) is:

$$(58) \quad C^1 = (1 - r)F(H, N, l) + w(1 - l) - (1 - \beta)pN$$

Maximizing C^1 with respect to the tenant's choice variables H , N , and l we get (for an interior maximum)

$$(59) \quad (1 - r)F_1 = 0$$

$$(60) \quad (1 - r)F_2 = (1 - \beta)p$$

$$(61) \quad (1 - r)F_3 = w$$

where F_i denotes the partial derivative of F with respect to its i th argument.

Let us denote by $|J|$ the determinant of the Jacobian matrix

$$J = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix}$$

where F_{ij} denotes the partial derivative of F_i with respect to its j th argument. By assumption $F_{ij} = F_{ji}$ and concavity of F implies $|J| < 0$, $F_{ii} < 0$ $i = 1, 2, 3$ and the determinants

$$\begin{vmatrix} F_{ii} & F_{ij} \\ F_{ji} & F_{jj} \end{vmatrix} > 0 \quad \text{for } i, j = 1, 2, 3, i \neq j$$

Let us further assume that $F_{ij} \geq 0$ for $i \neq j$. In other words, let us suppose that the marginal product of any input is a non-decreasing function of each of the other two inputs.

Using (59)–(61) it is easy to derive (62)–(67). These equations yield the ex-

$$(62) \quad H_r \equiv \frac{\partial H}{\partial r} = \frac{1}{(1 - r)^2 |J|} [(F_{13}F_{32} - F_{12}F_{33})(1 - \alpha)p + (F_{12}F_{23} - F_{13}F_{22})w] < 0$$

$$(63) \quad N_r \equiv \frac{\partial N}{\partial r} = \frac{1}{(1 - r)^2 |J|} [(F_{11}F_{33} - F_{13}^2)(1 - \alpha)p + (F_{21}F_{13} - F_{11}F_{23})w] < 0$$

$$(64) \quad l_r \equiv \frac{\partial l}{\partial r} = \frac{1}{(1 - r)^2 |J|} [(F_{12}F_{31} - F_{11}F_{32})(1 - \alpha)p + (F_{11}F_{22} - F_{12}^2)w] < 0$$

$$(65) \quad H_\beta \equiv \frac{\partial H}{\partial \beta} = \frac{-p(F_{13}F_{32} - F_{12}F_{33})}{(1 - r)|J|} > 0$$

$$(66) \quad N_\beta \equiv \frac{\partial N}{\partial \beta} = \frac{-p(F_{11}F_{33} - F_{13}^2)}{(1 - r)|J|} > 0$$

$$(67) \quad l_\beta \equiv \frac{\partial l}{\partial \beta} = \frac{-p(F_{12}F_{31} - F_{11}F_{32})}{(1 - r)|J|} > 0$$

pected results that (i) given β , the landlord's share in fertilizer costs (as well as w and p), the tenant will rent *less* land, put in *less* labor and buy *less* fertilizers the *larger* the value of r , the landlord's share in output, and (ii) given r , the tenant will rent *more* land, put in *more* labor and buy *more* fertilizers the *larger* the value of β .

Let us now turn to the landlord. Let us assume that he has one unit of land which he can either cultivate with hired labor or rent out to a sharecropper or do both in any desired proportion. Let q denote the proportion rented out. Let $G(1-q, x, y)$ denote the production function applicable to the land cultivated by the landlord with the help of hired labor, where x and y , respectively, denote the quantities of fertilizer and hired labor used. G is assumed to have the same properties as F . Let us assume that the landlord treats as given and independent of his own behavior the amount of fertilizers \bar{N} and labor L put in by the tenant. With these assumptions, the landlord's income (consumption reduces) to:

$$(68) \quad C^2 = G(1-q, x, y) - px - wy \\ + rF(q, \bar{N}, L) - \beta p\bar{N}$$

Maximizing C^2 with respect to choice variables q , x , and y we get (for an interior maximum):

$$(69) \quad -G_1 + rF_1 = 0$$

$$(70) \quad G_2 - p = 0$$

$$(71) \quad G_3 - w = 0$$

Let us denote by $|J'|$ the determinant of the Jacobian matrix

$$J' = \begin{bmatrix} G_{11} + rF_{11} & -G_{12} & -G_{13} \\ -G_{21} & G_{22} & G_{23} \\ -G_{31} & G_{32} & G_{33} \end{bmatrix}$$

Concavity of F and G implies $|J'| < 0$,

$$\begin{vmatrix} G_{ii} & G_{ij} \\ G_{ji} & G_{jj} \end{vmatrix} > 0 \quad \text{for } i \neq j, F_{ii}, G_{ii} < 0$$

As in the case of F , $G_{ij} > 0$ for $i \neq j$. It is easy to show that

$$(72) \quad q_r = \frac{\partial q}{\partial r} = \frac{-F_1(G_{22}G_{33} - G_{23}^2)}{|J'|} \geq 0$$

$$(73) \quad q_N = \frac{\partial q}{\partial \bar{N}} = \frac{-rF_{12}(G_{22}G_{33} - G_{23}^2)}{|J'|} \geq 0$$

$$(74) \quad q_L = \frac{\partial q}{\partial L} = \frac{-rF_{13}(G_{22}G_{33} - G_{23}^2)}{|J'|} \geq 0$$

It can be verified that x_r and y_r are non-positive and $x_{\bar{N}}$, $y_{\bar{N}}$, x_L , y_L are all non-negative. These are expected results: The landlord will rent out more land if his share in output of tenanted land is higher or if the tenant increase his inputs of fertilizers and labor. Also in each of these situations, the amount of land cultivated by the landlord himself will be correspondingly less. So will be the amounts of labor hired by him and the fertilizers purchased.

It should be noted that the landlord's choice variables do not depend directly on β , his share in the cost of fertilizers used by the tenant. This follows from our assumption that the landlord behaves as if he has no influence over the amounts of labor and fertilizer inputs used by the tenant.

Let us now turn to the equilibrium conditions:

$$(75) \quad H = q$$

$$(76) \quad N = \bar{N}$$

$$(77) \quad l = L$$

We have thus nine equations, (59)–(61), (69)–(71), and (75)–(77) to determine ten unknowns, H , N , l , q , \bar{N} , L , x , y , r , and β . Thus we have one unknown more than the number of equations and except in special cases, there will be more than one solution to the unknowns. In principle one can

$$(78) \quad \frac{dr}{d\beta} = \frac{(1-r)p(F_{13}F_{32} - F_{12}F_{33})}{[(F_{13}F_{32} - F_{12}F_{33})(1-\beta)p + (F_{12}F_{23} - F_{13}F_{22})w]} = \frac{-H_\beta}{H_r} > 0$$

solve the system for nine of the unknowns as functions of the tenth. Mathematics apart, there is one natural way of looking at the problem; i.e., as a market for land leases with the landlords as suppliers and tenants as demanders. There are two price-like variables; r the landlord's share of output and β the landlord's share in fertilizer cost. However, a single market-clearing condition that amount of land rented out by the landlords equals that leased in by the tenants can only determine one of these price variables as a function of the other. We treat r as a function of β and examine the behavior of r as β varies.

It can be shown that:

$$\frac{dr}{d\beta} = - \frac{q_l \cdot l_\beta + q_N N_\beta - H_\beta}{q_l \cdot l_r + q_N N_r - H_r + q_r}$$

After substituting the values of q_l , q_N , etc. and noting that $q_r=0$ in equilibrium, we get equation (78) which shows that $dr/d\beta$ is positive. This fact implies that the larger is the share of costs borne by landlords, the larger is the equilibrium rental share of crop. This rather straightforward result is worth keeping in mind, particularly because in the standard discussion on cost-sharing (see, for example, Adams and Rask) it is ignored by the assumption of keeping the rental share fixed and varying the cost share. Our result provides a simple explanation of a common observation in Indian agriculture: when the landlord participates in the costs, the rental share he is paid is much higher than otherwise. For example, the *Farm Management Survey in West Bengal* points to the *Bhagchasi* system in which the landlord does not share in the costs and gets about half of the crop produced

by the sharecropper whereas in the alternative *Krishani* system, the landlord himself covers most of the non-labor costs and usually gets about two-thirds of the crop raised by the sharecropper. Our result also implies that a government cannot try to implement rent-regulating legislation and at the same time also expect to induce the landlords to share more in tenant's fertilizer costs through sheer exhortations.

Equation (78) has other interesting implications. It can be rewritten as $H_r dr/d\beta + H_\beta = dH/d\beta = 0$. This means that regardless of the value of β in the relevant range (0, 1), the equilibrium share of the landlord in output will so adjust that the equilibrium amount of land leased out by the landlord (and leased in by the tenant) remains invariant.

Let us now consider a special case of the above model in which the tenant does not render wage labor and devotes all his working hours to the sharecropped land. For this case it can be shown that $dr/d\beta = 1-r/1-\beta$ which can be integrated to yield $r = (1-\theta) + \theta\beta$ where θ is some constant. From the fact that even if cost-sharing were absent (i.e., $\beta=0$), the equilibrium cropshare r will be positive, it follows that $0 < \theta < 1$. It also follows that $r > \beta$ for all β except $\beta=1$ when $r=1$. However $r=1$ rules out any tenancy. Hence we can assert that for all relevant values of β , i.e., in the interval (0, 1), the equilibrium output share of the landlord *exceeds* his share in costs.¹¹

It can also be shown for this special case (and for it only), that $dC^1/d\beta < 0$,

¹¹ We could prove this result only for the special case. In the general case it follows from (78) that $dr/d\beta < (1-r)/(1-\beta)$ and hence $r < (1-\theta) + \theta\beta$, with $0 < \theta < 1$. However this does not preclude $r \leq \beta$.

which implies that if the tenant had the choice of β , he would choose β to be zero. In other words, the tenant would prefer the landlord *not sharing* in costs. This apparently paradoxical result is explained as follows: we have seen that $dr/d\beta > 0$ implying that the larger the values of β the larger is the equilibrium value of r . Apparently the effect of lower r more than offsets the effect of lower β so that the optimum value of β from the tenant's point of view is zero. It can be shown that the landlord would prefer $\beta=1$ so that $r=1$. But this would rule out tenancy altogether. If, however, the social welfare indicator is the sum of the consumption of landlord and tenant, this sum can be shown to be invariant with respect to β so that the optimum β from the social point of view is any β in the interval $(0, 1)$.

VI. Uncertainty and Other Matters

In all of the analyses above we have ignored the intimate relationship of crop-sharing arrangements with problems of coping with uncertainty. In this paper we do not intend to deal with this important but very difficult issue in any detail. Let us just indicate here a possible theoretical way of approaching it, although we do not have many results to report.

Let us suppose A is a parameter representing production uncertainty (say, due to fluctuations of weather) and suppose equation (1) of Section I is rewritten (ignoring leisure) as

$$(79) \quad C^1 = (1-r)AF(H, l) + w(1-l)$$

and equation (8) as

$$(80) \quad C^2 = AG(1-q, 1+x) - wx + rAF(q, L)$$

Let us take a simple characterization of the uncertainty parameter A such that

$$(81) \quad A = \alpha u + \beta$$

where u is the random variate.

With E as the expectation operator and V for variance,

$$E(A) = \alpha E(u) + \beta \quad \text{and} \quad V(A) = \alpha^2 V(u)$$

For capturing the effects of parametric shifts in uncertainty we can vary the variance of A keeping the mean value of A constant by varying α and β in such a way that $d\beta/d\alpha = -E(u)$. Thus

$$(82) \quad \frac{dA}{d\alpha} = u + \frac{d\beta}{d\alpha} = u - E(u)$$

It is to be noted that with $\alpha > 0$, $(u - E(u))$ has the same sign as $(A - E(A))$.

We shall assume that both the tenant and the landlord maximize the expected value of their respective utility functions.

The necessary conditions for tenant's interior maximization are:

$$(83) \quad F_1 = 0$$

$$(84) \quad E[U'(C^1)\{(1-r)AF_2 - w\}] = 0$$

Similarly, the necessary conditions for the landlord's interior maximum are:

$$(85) \quad rF_1 - G_1 = 0$$

$$(86) \quad E[U'(C^2)\{AG_2 - w\}] = 0$$

Let us now analyze the equations (83)-(86). The Jacobian matrix $[a_{ij}]$ of equations (83) and (84) has the following elements:

$$a_{11} = F_{11}; \quad a_{12} = F_{12};$$

$$a_{21} = E[U'(C^1)(1-r)AF_{21}]$$

$$a_{22} = E[U'(C^1)(1-r)AF_{22}] \\ + E[U''(C^1)\{(1-r)AF_2 - w\}]$$

It is easy to check that the Jacobian is positive. From (82), (83), and (84), we may derive equation (87). The non-zero element in the column vector on the right-hand side of (87) can be shown to be positive¹² if $e^1 = -U''(C^1) \cdot C^1/U'(C^1)$, the

¹² Under the assumption of constant e^1

$$\phi(A) = [U'(C^1)F_2 + U''(C^1)\{(1-r)AF_2 - w\}F]$$

$$(87) \quad [a_{ij}] \begin{bmatrix} \frac{dH}{d\alpha} \\ \frac{dl}{d\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ -(1-r)E[\{U''(C^1)F_2 + U''(C^1)((1-r)AF_2 - w)F\}(u - Eu)] \end{bmatrix}$$

$$(88) \quad [b_{ij}] \begin{bmatrix} \frac{dq}{d\alpha} \\ \frac{dx}{d\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ -E[\{U''(C^2)G_2 + U''(C^2)(AG_2 - w)(G + rF)\}(u - E(u))] \end{bmatrix}$$

index of relative risk aversion, is assumed to be constant. This means that $dH/d\alpha < 0$; in other words, other things remaining the same, a larger importance of production uncertainty implies that the tenant will lease in *less* land under cropsharing.

Similarly, on the landlord side let us take equations (85) and (86). The Jacobian matrix $[b_{ij}]$ has the following elements.

$$\begin{aligned} b_{11} &= G_{11} + rF_{11}; \quad b_{12} = -G_{12}; \\ b_{21} &= -E[U''(C^2)AG_{21}]; \\ b_{22} &= E[U''(C^2)AG_{22}] \\ &\quad + E[U''(C^2) + \{AG_2 - w\}^2] \end{aligned}$$

It is easy to check that the Jacobian in this case also is positive. From (85) and (86), we obtain equation (88). Using the same technique as in footnote 12 we can prove that the non-zero element in the column vector on the right-hand side of (88) is positive if $e^2 = -U''(C^2) \cdot C^2 / U'(C^2)$, the index of relative risk aversion on the part of landlords, is assumed to be constant. This means that $dq/d\alpha > 0$; in other words, given other things (particularly r , w , and L), a larger importance of production uncertainty implies that the

landlord will tend to lease out *more* land to sharecroppers.

Because of the complicated calculations involved we have not been able to go any further. In any case, it seems that in our model increase in production uncertainty induces the tenant and the landlord to go in opposite directions in the market for land leases. We have not been able to find out the direction in the movement of the *equilibrium* proportion of land under sharecropping. Obviously much more work needs to be done in this area¹³ before even elementary comparative-static propositions can be made.

We may also point out here that apart from production uncertainty, other forms of uncertainty will also seriously affect the equilibrium percentage of area under sharecropping. For example, in our analysis above, increased production uncertainty induces the tenant to devote less labor to the tenanted farm and more to wage labor (this is largely because work

is a declining function of A . Hence

$$\begin{aligned} \text{and} \quad \phi(A) &> \phi(EA) && \text{when } A < EA \\ \phi(A) &< \phi(EA) && \text{when } A > EA \end{aligned}$$

so

$$E\phi(A)(-EA) < \phi(EA) \cdot E(A - EA) = 0$$

¹³ For not very rigorous treatments of this problem, see Cheung and Rao. Among other things, Cheung expects a larger incidence of share-tenancy in areas with higher production uncertainty and he mentions evidence from Chinese agriculture. In India, irrigation, apart from facilitating technical progress, has also a protective role against bad crop weather (there is some evidence of a negative correlation between percentage of area irrigated and variance of output) and yet in our empirical analysis there is a *positive* correlation between importance of irrigation and that of share-tenancy.

outside brings remuneration in the form of certain wage income whereas work on leased-in land involves uncertain income); but in a more realistic analysis, uncertainty of getting employment outside often drives the tenant to prefer a crop-sharing arrangement through which he shares the production uncertainty with the landlord. Similarly, on the landlord side there are the uncertainties of possible default of rent by the tenant in bad years and of availability of wage labor in peak seasons.

To sum up, in this paper we have tried to identify some of the economic factors that may contribute towards explaining regional variations in the incidence of cropsharing tenancy and to test for their significance on the basis of Indian State-level and village-level data. In the process we have built a simple theoretical model for analyzing sharecropping which may be useful in any further discussion on this form of tenancy. We have also tried to throw some theoretical light on the problem of cost-sharing by landlords.

Needless to say, our model here is very crude and it fails to capture many of the economic and institutional aspects of cropsharing tenancy as it is observed in peasant economies. In particular, we have hardly scratched the surface of any satisfactory analysis of various kinds of uncertainties that are relevant to cropsharing arrangements. We have also ignored the varying degrees of coexistence of crop-sharing with other forms of tenancy, like

fixed rent, on account of differences in uncertainty, the problems connected with the duration of lease contracts, and the various ways in which land-market imperfections (landlord dominance on the one hand and protective tenancy and rent-regulating legislation along with ways of evading them on the other) distort the simple rules of the competitive game we have assumed in most of this paper. Yet we believe that our basic framework will be useful in further explorations on this subject.

REFERENCES

- D. W. Adams and N. Rask, "Economics of Cost-Sharing Leases in Less Developed Countries," *Amer. J. Agr. Econ.*, Nov. 1968, 50, 935-42.
- S. N. S. Cheung, "Private Property Rights and Share-Cropping," *J. Polit. Econ.*, Nov.-Dec. 1968, 76, 1107-22.
- , "Transaction Costs, Risk Aversion and the Choice of Contractual Arrangements," *J. Law Econ.*, Apr. 1969, 12, 23-42.
- R. H. Day, "The Economics of Technological Change and the Demise of the Share-Cropper," *Amer. Econ. Rev.*, June 1967, 57, 427-49.
- D. G. Johnson, "Resource Allocation Under Share Contracts," *J. Polit. Econ.*, Apr. 1950, 58, 111-23.
- C. H. H. Rao, "Entrepreneurship, Management, and Farm Tenure Systems," unpublished.
- Government of India, Ministry of Food and Agriculture, *Farm Management Survey in West Bengal*.

On the Theory of the Competitive Firm Under Price Uncertainty

By AGNAR SANDMO*

In recent years several contributions have been made to the theory of the firm under uncertainty, removing the assumption that the demand for the product is known with certainty at the time when the output decision is made. In most of these papers the assumption is made that the objective of the firm is to maximize expected profits.¹ This is hardly a very satisfactory assumption, since it completely rules out risk averse behavior, and so many elementary facts of economic life seem to indicate a prevalence of risk aversion.

The present paper is intended as a systematic study of the theory of the competitive firm under price uncertainty and risk aversion. We assume that the decision on the volume of output to be produced must be taken prior to the sales date, at which the market price becomes known. The firm's beliefs about the sales price can be summarized in a subjective probability distribution. However, since the firm is unable to influence this distribution, the basic assumption that the firm is a price taker is retained—in a probabilistic sense.²

* Professor of economics, Norwegian School of Economics and Business Administration. This paper was written while I was a fellow of the Center for Operations Research and Econometrics, Université Catholique de Louvain. I am indebted to Jacques Drèze and Jean Jaskold Gabszewicz for their valuable comments.

¹ For some examples see the papers by Drèze and Gabszewicz, Kenneth Smith, Edward Zabel, and the book by Clement Tisdell.

² A similar approach is taken by Phoebus Dhrymes, Saul Hymans, John McCall, Bernt Stigum (1969a) and Hayne Leland (1969). Some interesting comments can also be found in Karl Borch (ch. 12, especially pp. 171–73).

It is perhaps most natural to interpret the model of the paper as being concerned with the short run. The firm makes its output decisions with sole regard for short-run profits and does not consider the relationship between this output policy and long-run policies for investment and finance. In a sense, it is a weakness of the model that it takes no account of this interrelatedness; but it may also be considered a strength, because a more complete model would make it necessary to draw up a much larger and more detailed list of assumptions about the economic environment of the firm than is needed for the present paper. The results presented here are thus compatible with several alternative sets of assumptions about investment opportunities, financial markets, and the structure of ownership. It is only essential to assume that short-run output decisions are dominated by a concern for short-run profits.

Occasionally, especially in Section III, we shall also find it convenient to use the model to analyze some long-run problems. It then becomes necessary to assume that these long-run elements have implicitly been accounted for. This is hardly satisfactory. Still, it is a useful simplification with long traditions in the theory of the firm.

We shall assume that the firm's attitude towards risk can be summarized by a von Neumann-Morgenstern utility function. This may be a strong assumption, because in many firms decisions are typically taken by a group of individuals, and group preferences may not always satisfy the

transitivity axiom required for the existence of a utility function. It is therefore possible that this approach implicitly assumes that the firm's reactions to changes in its environment are more predictable and stable than they really are. However, there are still many firms in which decisions are essentially made by one person, and there are presumably firms in which preferences are sufficiently similar within the group of decision makers to guarantee the existence of a group preference function. This provides justification for the approach taken in this paper.

I. Optimal Output under Uncertainty

We assume that the objective of the firm is to maximize the expected utility of profits. The utility function of the firm is a concave, continuous and differentiable function of profits, so that

$$(1) \quad U'(\pi) > 0, \quad U''(\pi) < 0$$

Thus, the firm is assumed to be risk averse. It is well known that in order for a utility function to satisfy the von Neumann-Morgenstern axioms without giving rise to St. Petersburg phenomena, it must be bounded from above.³ Strictly speaking, then, equation (1) holds only in the range below the upper bound of U .

The cost function of the firm is

$$(2) \quad F(x) = C(x) + B,$$

where x is output, $C(x)$ is the variable cost function, and B is "fixed cost." About the variable cost function we make the following general assumptions:

$$(3) \quad C(0) = 0, \quad C'(x) > 0$$

The firm's profit function can now be defined as

$$(4) \quad \pi(x) = px - C(x) - B,$$

where p is the price of output, assumed to

be a (subjectively) random variable with density function $f(p)$ and expected value $E[p] = \mu$. Naturally, p is restricted to be nonnegative. This means that, once x has been chosen, the firm's maximum loss is $(-C(x) - B)$. Clearly also, $\pi(0) = -B$.

The expected utility of profits can be written as

$$E[U(px - C(x) - B)],$$

where E is the expectations operator. Differentiating with respect to x , we obtain as necessary and sufficient conditions for a maximum:

$$(5) \quad E[U'(\pi)(p - C'(x))] = 0,$$

$$(6) \quad D = E[U''(\pi)(p - C'(x))^2 - U'(\pi)C''(x)] < 0$$

It is interesting to note that in order for the second-order condition (6) to hold, it is not necessary to assume increasing marginal cost.

For the remainder of Section I and in Section II, we assume that (5) and (6) determine a non-zero, finite and unique solution to the maximization problem. The problems of existence and of corner solutions will be discussed in Section III.

One question which is naturally raised by the introduction of price uncertainty is this: how does the optimal output compare with the well-known competitive solution under certainty? Under certainty, the solution is characterized by equality between price and marginal cost. There is no obvious way of making such a comparison, but one possible and appealing specification of the problem is this: what is the optimal output under uncertainty as compared with the situation where the price is known to be equal to the expected value of the original distribution? Referring to the latter level of output as the certainty output, we shall now show that *under price uncertainty, output is smaller than the*

³ See on this point Kenneth Arrow, who also argues that U must be bounded from below.

certainty output. This is a generalization of a theorem of McCall, who proves a similar result for the case of a utility function with constant absolute risk aversion.

The first-order condition (5) can be written as

$$(7) \quad E[U'(\pi)p] = E[U'(\pi)C'(x)]$$

Subtract $E[U'(\pi)\mu]$ on each side of this equation. We then get

$$(8) \quad E[U'(\pi)(p-\mu)] = E[U'(\pi)(C'(x)-\mu)]$$

Since $E[\pi] = \mu x - C(x) - B$ (from the definition of profits), we have that $\pi = E[\pi] + (p-\mu)x$. Clearly

$$(9) \quad U'(\pi) \leq U'(E[\pi]) \quad \text{if } p \geq \mu$$

It follows immediately that

$$(10) \quad U'(\pi)(p-\mu) \leq U'(E[\pi])(p-\mu)$$

This inequality holds for all p . For if $p \leq \mu$, the inequality sign in (9) is reversed, but then multiplication by $(p-\mu)$ will still make \leq hold in (10). Taking expectations on both sides of (10) and noting that $U'(E[\pi])$ is a given number, we obtain

$$E[U'(\pi)(p-\mu)] \leq U'(E[\pi])E[p-\mu]$$

But, here the right-hand side is equal to zero by definition, and so the left-hand side is negative. Then we know that the right-hand side of (8) is negative also. But this can be written as

$$E[U'(\pi)](C'(x) - \mu) \leq 0,$$

and, since marginal utility is always positive, this implies

$$(11) \quad C'(x) \leq \mu$$

That is, optimal output is characterized by marginal cost being less than the expected price. Now under certainty the only types of cost curves compatible with competitive assumptions are those for which the marginal cost curve is either everywhere increasing or else U-shaped. In those cases,

(11) proves our statement above. Equation (11) is, of course, also valid for constant or decreasing marginal cost, but then the competitive output is not well defined.

This result is not the only conceivable answer to the question of the effect of uncertainty on the output decision. Following Jacques Drèze and Franco Modigliani, we may describe our result as concerned with the *overall* impact of uncertainty. However, one may also be interested in the question of the *marginal* impact; i.e., the effect of making a given distribution "slightly more risky." It is not obvious how this can be formalized; in the following we shall adopt a procedure used in Sandmo.

Let us define a small increase in risk as a "stretching" of the probability distribution around a constant mean. This requires the introduction of two shift parameters, one multiplicative and one additive. Thus, let us write price as

$$\gamma p + \theta,$$

where γ is the multiplicative shift parameter and θ is the additive one. An increase of γ alone (from the point $\gamma=1, \theta=0$) will "blow up" all values of p ; it will therefore increase the mean as well as the variance. To restore the mean we have to reduce θ simultaneously, so that

$$dE[\gamma p + \theta] = 0, \quad \text{or} \quad \mu d\gamma + d\theta = 0, \quad \text{i.e.,}$$

$$(12) \quad \frac{d\theta}{d\gamma} = -\mu$$

We can now write the profit function as $\pi(x) = (\gamma p + \theta)x - C(x) - B$ and differentiate with respect to γ , taking account of (12). The result is then

$$(13) \quad \frac{\partial \pi}{\partial \gamma} = -x \cdot \frac{1}{D} E[U''(\pi)(p-\mu)(p-C'(x))] - \frac{1}{D} E[U'(\pi)(p-\mu)]$$

Of these two terms, the last one is clearly

negative (from the proof above and from the second-order condition). However, the sign of the first term is in general indeterminate, so that at the present level of generality it does not seem possible to make a precise statement about the marginal impact of uncertainty.

There is one special case in which we would expect the marginal impact of uncertainty to become identical to the overall impact. That is in the case where we start from the certainty of $p = \mu$ and replace this certain price by a probability distribution with all outcomes concentrated in the neighborhood of μ . This is not too easily handled, since our stretching procedure breaks down in that case. However, we can get around this difficulty by noting that, when price is known to be equal to μ , we must have $C'(x) = \mu$. Then the first term in (13) becomes

$$-x \cdot \frac{1}{D} E[U''(\pi)(p - \mu)^2],$$

which is certainly negative. Thus, both terms in (13) are negative, and their signs depend only on the assumption of risk aversion. The connection with the overall impact of uncertainty is thereby established.

II. The Comparative Statics of the Firm

Simply assuming the existence of risk aversion is a very weak restriction on the firm's attitudes to risk. Further restrictions on the utility function may be introduced by means of the Arrow-Pratt risk aversion functions:

$$\text{Absolute risk aversion: } R_A(\pi) = -\frac{U''(\pi)}{U'(\pi)}$$

$$\text{Relative risk aversion: } R_R(\pi) = -\frac{U''(\pi)\pi}{U'(\pi)}$$

It seems reasonable to assume that $R_A(\pi)$ is a decreasing function of π . This would reflect the hypothesis that as a

decision maker becomes wealthier (in terms of income, profit etc.), his risk premium for any risky prospect, defined as the difference between the mathematical expectation of the return from the prospect and its certainty equivalent, should decrease, or at least not increase. If $R_R(\pi)$ is increasing, this means that the elasticity of the risk premium with respect to π is less than one in absolute value. Arrow argues that there are good theoretical and empirical reasons for making this assumption, but the evidence for it does not seem conclusive, and we shall not commit ourselves to a specific hypothesis as to the form of $R_R(\pi)$.⁴

One of the basic results in the theory of the firm under certainty is that fixed costs do not matter in the sense that once a strictly positive output level has been chosen, this output is unaffected by an infinitesimal increase in fixed costs. This is not so under uncertainty. Differentiating in (5) with respect to B , we obtain

$$(14) \quad \frac{\partial x}{\partial B} = \frac{1}{D} E[U''(\pi)(p - C'(x))]$$

Decreasing absolute risk aversion is a necessary and sufficient condition for $\partial x / \partial B$ to be negative. The proof of this is as follows: Let $\bar{\pi}$ be the level of profits when $p = C'(x)$. Then, since $R_A(\pi)$ is decreasing⁵

⁴ Some remarks on the empirical evidence can be found in the article by Joseph Stiglitz. For derivations of the risk aversion functions the reader is referred to the contributions of Arrow and John Pratt. Hypotheses about the risk aversion functions have been applied to portfolio theory by Arrow, to insurance purchasing and to taxation and risk-taking by Jan Mossin (1968a, b), and to the analysis of saving decisions by Sandmo. Several other examples of application could easily be given.

⁵ This must be interpreted with care. We are interested in the properties of the risk aversion function at the optimum position, i.e., for the output level $x = x^*$ which is the solution to (5). For this given output level, (15) is certainly true. It is important to note that this *local* relationship is independent of the *global* lack of any one-to-one relationship between the algebraic signs of profits and marginal revenue.

$$(15) R_A(\pi) \leq R_A(\bar{\pi}) \text{ for } p - C'(x) \geq 0$$

Substituting from the definition of $R_A(\pi)$, we obtain

$$(16) -\frac{U''(\pi)}{U'(\pi)} \leq R_A(\bar{\pi}) \text{ for } p - C'(x) \geq 0$$

(Note that $R_A(\bar{\pi})$ is a given number and not a random variable.) We know of course that

$$(17) -U''(\pi)(p - C'(x)) \leq 0 \text{ for } p - C'(x) \geq 0,$$

since marginal utility is positive. Now multiply (16) by the left-hand side of (17). We then get

$$U''(\pi)(p - C'(x)) \geq -R_A(\bar{\pi})U'(\pi)(p - C'(x))$$

This holds for all p . For if $p \leq C'(x)$, the inequality in (16) is reversed, but so is that in (17). Now taking expected values we obtain

$$E[U''(\pi)(p - C'(x))] \geq -R_A(\bar{\pi})E[U'(\pi)(p - C'(x))]$$

But by the first-order condition (5), the right-hand side is equal to zero, and the left-hand side is accordingly positive. But then the derivative (14) is negative and our proposition is proved.

Is this conclusion in itself intuitively plausible? This question may perhaps best be judged by considering whether a lump sum tax or a lump sum subsidy would be the most appropriate policy measure for making the firm increase its output. Economic intuition seems strongly to suggest the latter alternative, which is exactly what our result implies.

We turn now to an examination of the firm's supply function. Since the price is seen by the firm as a random variable, it does not make sense to speak about the effect of an "increase in price." It seems natural, however, to discuss the closely related problem of an increase in the mathematical expectation of the price with higher central moments constant. We can

do this in the following way: Let us write price as $p + \theta$, where θ is again an additive shift parameter. Increasing θ is equivalent to moving the probability distribution to the right without changing its shape. Differentiating (5) with respect to θ and evaluating the derivative at $\theta = 0$ we obtain

$$\frac{\partial x}{\partial \theta} = -x \cdot \frac{1}{D} E[U''(p - C'(x))] - \frac{1}{D} E[U'(\pi)],$$

or, substituting from (14),

$$(18) \quad \frac{\partial x}{\partial \theta} = -x \frac{\partial x}{\partial B} - \frac{1}{D} E[U'(\pi)]$$

This expression is similar to the Slutsky equation familiar from demand analysis. It says that the firm's response to an increase in expected price can be decomposed into two separate effects, one of which is analogous to a decrease in fixed costs, and the other one is a pure substitution effect. Of the latter effect we can immediately say that it is positive. As for the sign of the former effect we can draw on our previous result to conclude that *decreasing absolute risk aversion is a sufficient condition for $\partial x / \partial \theta$ to be positive*, i.e., for an upward-sloping supply curve. Again the implication of decreasing absolute risk aversion seems intuitively plausible. It implies, e.g., that in order to increase output the government should consider a per unit subsidy, rather than a per unit tax, as the appropriate policy measure.⁶

Another well-established result in the theory of the firm is that a change in a proportional rate of profit taxation will have no effect on the level of output. A priori there is no reason to expect this result to hold under uncertainty.

⁶ The interested reader who wishes to see an example where the possibility of a downward-sloping supply curve does occur may consider the simple case of a quadratic utility function and constant marginal cost, where the supply curve bends backward for expected price sufficiently high.

With price uncertainty the question of loss offset provisions becomes important. If there is no loss offset, the profit function of the firm becomes

$$\pi(x) = \begin{cases} px - C(x) - B & \text{for } p \leq \frac{C(x) + B}{x} \\ (px - C(x) - B)(1 - t) & \text{for } p > \frac{C(x) + B}{x} \end{cases}$$

On the other hand, if there is full loss offset, the profit function can be written as

$$\pi(x) = (px - C(x) - B)(1 - t) \quad \text{for all } p$$

It is not easy to decide which of these two assumptions is the more interesting and realistic one. Full loss offset presupposes that the firm or its owner(s) has other income from which any loss can be deducted. In fact, tax laws in many countries do provide for loss offset, either against other income or against future profits, so that there may be reasons for concentrating attention on this case.⁷

With full loss offset expected utility is

$$E[U((px - C(x) - B)(1 - t))],$$

and the first-order condition becomes

$$(19) \quad E[U'(\pi)(p - C'(x))] = 0,$$

as before, since the multiplicative factor $(1 - t)$ can be factored out.

Differentiating in (19) with respect to t yields

$$(20) \quad \frac{\partial x}{\partial t} = \frac{1}{1 - t} \cdot \frac{1}{D} \cdot E[U''(\pi)\pi(p - C'(x))]$$

It can be shown that *increasing the tax rate will increase, leave constant or reduce output*

⁷ This argument is not entirely satisfactory, however. If "other income" or "future profits" are at least partially determined by the firm's own actions, they should presumably be integrated into the model.

according as relative risk aversion is increasing, constant, or decreasing.

If $R_R(\pi)$ is increasing, we must have that

$$(21) \quad -\frac{U''(\pi)\pi}{U'(\pi)} \geq R_R(\pi) \quad \text{for } p - C'(x) \geq 0$$

Multiplying this by $-U'(\pi)(p - C'(x))$ yields

$$U''(\pi)\pi(p - C'(x)) \leq -R_R(\pi)U'(\pi)(p - C'(x)),$$

and by the argument used in the proof above, this inequality holds for all p . Taking expectations, the right-hand side vanishes, and we have that

$$E[U''(\pi)\pi(p - C'(x))] \leq 0$$

From this it follows that $\partial x / \partial t$ is positive in the case of increasing relative risk aversion. The proof of the rest of the statement follows immediately.

III. Profits, Entry, and Returns to Scale

It is well known that under certainty increasing marginal cost is necessary for the existence of a competitive optimum for the firm. This is not so under uncertainty, as we shall now demonstrate.⁸

Consider first the case where marginal cost is constant. Then concavity and boundedness of U as a function of π is sufficient to show that there exists a finite $x = x^*$ which gives a maximum of U . The case $C''(x) > 0$ is equally simple, because increasing marginal cost only reinforces the concavity of U as a function of x . It follows also that the case of a U-shaped marginal cost curve is only slightly more complicated for then U will be concave in x in the region for which $C'(x) \geq \min C'(x)$.

Note also that in the case of decreasing MC followed by constant MC the above

⁸ For a rigorous discussion of the existence of optimal policies under uncertainty the reader is referred to Leland (1970).

argument remains valid; there will be a determinate optimal level of output for the firm. The troublesome case is where MC is everywhere decreasing and boundedness of the utility function no longer guarantees the existence of an optimal policy. However, it remains true that decreasing MC is not a sufficient condition for the nonexistence of an optimal output level; thus a market *may* be competitive even under this assumption.

So far, we have assumed the existence of an interior maximum for the firm; i.e., we have assumed that the optimal level of output is strictly positive. But we know from received theory that even if the condition "price=marginal cost" determines a local maximum of profits, the maximum need not, even if it is a unique interior maximum, give us the global maximum. The reason is simply that the interior maximum may result in negative profits, so that the best policy is to produce nothing at all. In other words, production will take place at a positive level if, and only if, the best positive production level results in nonnegative profit.

Let x^* be the output level which is the solution to (5) and satisfies (6). Then x^* will also give a global utility maximum, provided that

$$(22) \quad E[U(\mu x^* - C(x^*) - B)] \geq U(-B)$$

It will be recalled that $-B$ is the level of profit for $x=0$.⁹

Developing the left-hand side of (22) in a Taylor series around the point $p=\mu$ we obtain, neglecting higher-order terms,

$$\begin{aligned} E[U(\mu x^* - C(x^*) - B) + U'(\mu x^* - C(x^*) \\ - B)x^*(p - \mu) + \frac{1}{2}U''(\mu x^* - C(x^*) \\ - B)x^{*2}(p - \mu)^2] \geq U(-B) \end{aligned}$$

The second term on the left-hand side is zero by definition. Rearranging the re-

maining terms and dividing through by $U'(\mu x^* - C(x^*) - B)$ so as to make the expressions invariant under linear transformations of the utility function, we then get

$$\begin{aligned} (23) \quad & \frac{U(\mu x^* - C(x^*) - B) - U(-B)}{U'(\mu x^* - C(x^*) - B)} \\ & \geq -\frac{1}{2} \frac{U''(\mu x^* - C(x^*) - B)}{U'(\mu x^* - C(x^*) - B)} x^{*2} E[p - \mu]^2 \end{aligned}$$

Both sides of this inequality have the dimension of money. The factors on the right-hand side are the risk aversion function, evaluated at the expected level of profit for $x=x^*$, and the variance of sales, $x^{*2}E[p-\mu]^2$. Since both these factors are positive, the left-hand side must also be positive, and with a strictly increasing utility function this implies that

$$\mu x^* - C(x^*) - B > -B,$$

or

$$(24) \quad \mu > \frac{C(x^*)}{x^*},$$

i.e., at the optimum *expected price must be larger than average cost, so that the firm requires positive expected profit in order to choose a positive output level*. It should be stressed that "positive" here means "strictly positive." If expected profit for $x=x^*$ were zero, (23) would not be satisfied, and the output level of zero would be chosen. We conclude, therefore, that competitive equilibrium under price uncertainty and risk aversion requires the existence of positive profits.¹⁰

It is interesting to study the role of risk aversion in the long-run equilibrium posi-

⁹ The argument here could equally well be carried out under the "long-run" assumption that $B=0$.

¹⁰ As in any partial equilibrium analysis this statement is somewhat incomplete. Implicit in it is the assumption that by not producing anything the owners of firms can make a sure return by employing their resources elsewhere in the economy. If this return is strictly positive, "normal profits" should be included among the firms' costs.

tion.¹¹ We assume therefore, to make the discussion simpler, that firms have identical cost functions and identical probability beliefs. Looking at (23) it is easy to see that a ("almost") risk-neutral firm will require only a nonnegative profit to enter the industry; in other words, as long as any positive level of expected profit remains, risk-neutral firms will enter. It is also clear from (23) that firms with "very high" risk aversion will not enter the industry at all, or they will be marginal firms in the sense that a very small decrease in expected price will make them leave the market. The risk neutral firms will of course set marginal cost equal to expected price (assuming U-shaped cost curves), while the risk-averse firms in the industry will choose output levels for which marginal cost is less than expected price. In general, the distribution of output and expected profit among firms will vary with their degree of risk aversion. Expected profit will be highest for those firms which come very close to being risk neutral and have the highest output in the industry. This observation confirms a view which has long traditions in economic theory, viz. to regard profit as a reward to risk-bearing.

Let us now turn to the case where marginal cost is constant or decreasing. We have shown that this case is not inconsistent with competitive assumptions. However, if one or a few firms are much less risk averse than the others, they may choose very high output levels and thereby lower expected price so much that the others will leave the industry. An uneven distribution of risk aversion may therefore be a source of oligopolistic concentration in its own right.

IV. Concluding Remarks

There are many ways in which this

analysis can be extended and generalized. We have had nothing to say on the subject of the multiproduct firm, which is of particular interest under uncertainty, since the firm is able to spread its risks by output diversification.¹² Neither have we had anything to say about the role of inventories under demand uncertainty. Finally, investment and financing decisions can hardly be given adequate treatment in the present framework.

It would also be interesting to place the competitive firm facing price uncertainty in a general equilibrium framework. This would require a different type of analysis from that of Debreu, in which there exists a complete set of markets for contingent commodities and the firm bears no risk at all. An alternative approach is contained in a recent paper by Stigum (1969b), in which firms do bear risks and entrepreneurs display risk averse behavior. Evidently, alternative models can be constructed with different assumptions about ownership and market opportunities: the theory of the firm developed in the present paper presumably will fit into some, but not all, of these models.

REFERENCES

- K. J. Arrow, *Aspects of the Theory of Risk-Bearing*, Helsinki 1965.
- K. H. Borch, *The Economics of Uncertainty*, Princeton 1968.
- G. Debreu, *Theory of Value*, New York 1959.
- P. J. Dhrymes, "On the Theory of the Monopolistic Multiproduct Firm under Uncertainty," *Int. Econ. Rev.*, Sept. 1964, 5, 239-57.
- J. Drèze and J. J. Gabszewicz, "Demand Fluctuations, Capacity Utilization and Prices," *Operations Research Verfahren*, 1967, 3, 119-41.
- J. Drèze and F. Modigliani, "Consumption Decisions under Uncertainty," CORE Dis-

¹¹ For the following discussion, which is essentially long-run, it is appropriate to assume $B=0$; in the long run all costs are variable costs.

¹² This problem has been studied by Dhrymes for the special case of a quadratic utility function.

- cussion Paper No. 6906, Louvain 1969.
- S. H. Hyman, "The Price-Taker: Uncertainty, Utility and the Supply Function," *Int. Econ. Rev.*, Sept. 1966, 7, 346-56.
- H. E. Leland, "The Theory of the Firm Facing Uncertain Demand," mimeo. Stanford Univ., 1969.
- , "On the Existence of Optimal Policies under Uncertainty," mimeo. Stanford Univ., 1970.
- J. J. McCall, "Competitive Production for Constant Risk Utility Functions," *Rev. Econ. Stud.*, Oct. 1967, 34, 417-20.
- J. Mossin (1968a), "Aspects of Rational Insurance Purchasing," *J. Polit. Econ.*, July/Aug. 1968, 76, 553-568.
- (1968b), "Taxation and Risk-Taking: An Expected Utility Approach," *Econometrica*, Feb. 1968, 35, 74-82.
- J. W. Pratt, "Risk Aversion in the Small and in the Large," *Econometrica*, Jan./Apr. 1964, 32, 122-136.
- A. Sandmo, "The Effect of Uncertainty on Saving Decisions," *Rev. Econ. Stud.*, July 1970, 37, 353-60.
- K. R. Smith, "The Effect of Uncertainty on Monopoly Price, Capital Stock and Utilization of Capital," *J. Econ. Theory*, June 1969, 1, 48-59.
- J. E. Stiglitz, "The Effects of Income, Wealth and Capital Gains Taxation on Risk-Taking," *Quart. J. Econ.*, May 1969, 83, 263-83.
- B. P. Stigum (1969a), "Entrepreneurial Choice over Time under Conditions of Uncertainty," *Int. Econ. Rev.*, Oct. 1969, 10, 426-42.
- (1969b), "Competitive Equilibria under Uncertainty," *Quart. J. Econ.*, Nov. 1969, 83, 533-61.
- C. A. Tisdell, *The Theory of Price Uncertainty, Production and Profit*, Princeton 1969.
- E. Zabel, "A Dynamic Model of the Competitive Firm," *Int. Econ. Rev.*, June 1967, 8, 194-208.

The Effect of Tariffs on Production, Consumption, and Trade: A Revised Analysis

By J. CLARK LEITH*

The influence of tariffs on production, consumption, and trade has long occupied a significant place in the literature of international economics. Until recently these influences were analyzed entirely in the context of tariffs on final goods, but the introduction of tariffs on intermediate goods into the discussion has resulted in a renewed interest and revision of the analysis. The new approach has been carried forward in both general and partial equilibrium models, with major interest centering on the partial equilibrium aspects of the stimulus to production in the form of the theory of "effective protection."¹ In this paper, we bring together the production effect of tariffs with the consumption and use effects to focus on the net effect of a set of tariffs on imports of a commodity. The purpose is to show that the assumed elasticity of substitution between inputs in production is a significant ele-

ment in the partial equilibrium² analysis. We specify two alternative models which differ in the elasticity of substitution. In the first, we develop a model under the assumption of a zero elasticity of substitution commonly found in much of the effective protection literature. In the second, we consider the consequences of a model in which there is a unitary elasticity of substitution between inputs in production.

I. Assumptions

Consider the market under free trade for a particular good, J , which is one of many. Assume that foreign produced J is a perfect substitute for the domestically produced J , and that the foreign supply is infinitely elastic at a price that is below the autarky price but not below the zero offer price of domestic production. Under free trade, domestic producers offer that amount at which the world price equals their marginal costs, consumers and users purchase that amount at which the world price equals the marginal usefulness of J to them, and the difference is made up of imports.

If we now impose a set of tariffs affecting the market for good J , the markets for J 's inputs, and the markets where J is used as an input, then domestic production

* Assistant professor, department of economics, University of Western Ontario, and currently visiting at the University of Ghana under a twinning arrangement between the two departments. He is indebted to W. M. Corden, Herbert Grubel, Peter J. Lloyd, James R. Melvin, P. O'Brien, J. R. Williams, and G. D. Wood for helpful comments on an early draft. None, however, should be held responsible for what is presented here.

¹ The seminal partial equilibrium contributions are by W. M. Corden (1966) and Harry Johnson (1965). Corden also makes a beginning in the task of extending the partial equilibrium analysis in the direction of a general equilibrium system. And, there are the important early contributions to general equilibrium analysis of Ronald McKinnon and William Travis (1964). More recent general equilibrium contributions include Roy Ruffin, and V. K. Ramaswami and T. N. Srinivasan.

² The analysis is restricted to partial equilibrium for a number of general equilibrium effects are ignored, including cross-elasticities of demand effects, income effects, the effects of changes in factor prices on other industries, and the effects of tariffs on the balance of trade and the exchange rate.

consumption, use, and hence imports will be changed. In analyzing these changes, assume further that there is domestic production and trade both with and without tariffs, and that there is a production function homogeneous of degree one for each good that relates inputs of tradeable materials and a primary factor to output.³

II. Zero Elasticity of Substitution Case⁴

First, consider the effect of tariffs on production of J when we assume a zero elasticity of substitution between all inputs used to produce it. The change in production of J is⁵

$$(1) \quad dS_j = \left(t_j - \sum_i a_{ij} t_i \right) e_j S_j$$

where I is the typical input, and the a_{ij} coefficient is the value of I used per dollar of J output, at free trade prices. The term $\sum_i a_{ij} t_i$ describes the upward shift in the supply curve due to the tariffs on inputs, and the difference between that and the price increase of the output ($t_j - \sum_i a_{ij} t_i$) is the *net* rate of protection received by producers of good J .⁶ It is the net protec-

tion, together with the elasticity of supply of the original supply curve e_j , that describes the expansion of domestic production. Equation (1) can be simplified by replacing $(t_j - \sum_i a_{ij} t_i)$ with the single symbol π_j representing the net rate of protection afforded domestic output by the tariff structure:

$$(2) \quad dS_j = \pi_j e_j S_j$$

The change in household consumption is

$$(3) \quad dC_j = \eta_j t_j C_j$$

and the change in intermediate use is

$$(4) \quad dU_j = \sum_i a_{ji} \left(t_i - \sum_j a_{ij} t_j \right) e_i S_i \\ = \sum_i a_{ji} \pi_i e_i S_i$$

which describes the shift of the demand curve, for $\pi_i e_i S_i$ is the expanded production in the typical using industry, I , and the a_{ji} term indicates the impact this has on industry J .

The change in imports is the sum of (3) and (4) minus (2).

$$(5) \quad dM_j = -\pi_j e_j S_j + \eta_j t_j C_j + \sum_i a_{ji} \pi_i e_i S_i$$

or, to be explained below

$$dM_j = -f_j e_j S_j + \eta_j t_j C_j + \sum_i a_{ji} f_i e_i S_i$$

Our attention so far has been confined to influence of tariffs on the market for the typical *product*. We can now indicate how this analysis corresponds to the idea embodied in the effective protection concept of protection of a *process*. The effective rate of protection is the proportionate change in the price of the primary factor that is made possible by the imposition of tariffs (f_j). The formula is

$$(6) \quad f_j = \frac{t_j - \sum_i a_{ij} t_i}{1 - \sum_i a_{ij}} = \frac{\pi_j}{v_j},$$

³ The assumption of a single primary factor is clearly an oversimplification. However, without being excessively arbitrary, it can be taken to represent a composite of primary factors, and changes in its price as the weighted average change in the price of the primary factors. See Lloyd for an approach that uses multiple primary inputs.

⁴ This section draws on the analysis developed by Rachel Dardis, and Johnson (1969). The latter, in turn, has been used by Bela Balassa (1965, 1967), and Leith and G. L. Reuber in discussing the restriction of imports due to tariffs. While it could be argued that the shifts of the demand and supply curves described here are not partial equilibrium effects in the sense that the influence of more than a change in the price of J is considered, we do confine our attention to the effects on this *one* market of *one set* of tariffs. We ignore effects such as changes in relative prices of other substitutes and complements on both the production and consumption sides.

⁵ Throughout the paper we will utilize linear approximations for discrete changes along supply and demand curves.

⁶ The concept of the net protection received by producers was used by Paul and Ronald J. Wonnacott.

where v_j is value-added per dollar of output of industry J . This is clearly related to the production effect indicated in equations (1) and (2) above, and it can be shown that the effective rate of protection together with the elasticity of supply of the primary factor is an equivalent way of representing the production effect. Let S_{fj} represent the quantity of the primary factor used in production of J , ϵ_j the elasticity of supply of the primary factor used in the production of J , and α_{fj} is the quantity of the primary factor that is used per unit of output in the free trade equilibrium situation. Thus, $\alpha_{fj} = S_{fj}/S_j$, and under free trade

$$(7) \quad S_j = S_{fj}/\alpha_{fj},$$

and with protection⁷

$$(8) \quad S_j(1 + dS_j/S_j) = \frac{S_{fj}(1 + dS_{fj}/S_{fj})}{\alpha_{fj}(1 + d\alpha_{fj}/\alpha_{fj})}$$

Because of zero substitution, $d\alpha_{fj}/\alpha_{fj} = 0$. Now define

$$\epsilon_j \equiv \frac{dS_{fj}/S_{fj}}{f_j}$$

Therefore (8) may be written

$$(9) \quad S_j(1 + dS_j/S_j) = \frac{S_{fj}}{\alpha_{fj}}(1 + \epsilon_j f_j),$$

but we may eliminate the initial conditions (7) to obtain

$$(9') \quad dS_j = \epsilon_j f_j S_j$$

Equation (9') is thus equivalent to equation (2) as a way of representing the expansion of domestic output.

Although either the approach of equation (9') or that of equation (2) tells us by how much output changes, we do not have this choice of approaches when we consider the influence of tariffs on consumption and

imports. Consumption and trade can be measured only in units of the product, not units of the process.

Consider also the pull of resources into the production of good J due to protection.⁸ The proportionate change in use by industry J of the typical input I (dS_{ij}/S_{ij}) is readily specified under the fixed coefficient assumption: the proportionate change in use of any input is equal to the proportionate change in output. Hence

$$(10) \quad \frac{dS_{ij}}{S_{ij}} = \frac{dS_j}{S_j} = \pi_j e_j,$$

where I is any input, material or primary. Again, the net rate of protection and the supply elasticity must be known. From this we can determine the proportionate change in the price of the typical input I (dp_{ij}/p_{ij}) induced by the expansion of output

$$(11) \quad dp_{ij}/p_{ij} = \pi_j(e_j/\epsilon_j)$$

where e_{ij} is the elasticity of supply of input I to industry J . Under the assumption of infinitely elastic supplies of tradeables, the proportionate change in the price of a tradeable input is clearly zero. However, if I is a non-traded material input, the change in its price is given by equation (11).⁹ The proportionate change in the price of the primary factor (i.e., the effective rate of protection), is a special case of (11):

$$(11') \quad dp_{ij}/p_{ij} = \pi_j(e_j/\epsilon_j) \\ \text{(for } I = \text{primary factor input),}$$

and since $\epsilon_j = v_j e_j$

$$(11'') \quad dp_{ij}/p_{ij} = \pi_j/v_j \\ \text{(for } I = \text{primary factor input).}$$

⁸ The concepts that follow here draw in part on Benton Massell.

⁹ We must assume, however, that there are no tariffs on inputs used in the production of non-traded inputs. (See Leith (1968), p. 591.)

⁷ Note that whether or not the quantity of the primary factor used per unit of output will change depends on whether there is substitution of inputs.

Finally, consider the effect of tariffs on domestic costs under protection. The stimulus to production is π_j (or f_j), but the higher marginal cost of domestic production is still given by the nominal tariff.¹⁰ This is made up of the higher cost of inputs due to tariffs on them, plus the induced higher prices of primary and non-traded material inputs, the induced portion being shared in proportion to the inputs' supply elasticities. This may be seen by rearranging (11) and summing over all inputs, yielding

$$(12) \quad t_j = \sum_i \left(a_{ij} t_i + \frac{dp_{ij}}{p_{ij}} \cdot \frac{e_{ij}}{e_j} \right) \\ \text{(for } I = \text{all inputs).}$$

III. Unitary Elasticity of Substitution Case

Consider now the effect of tariffs on production, consumption, use, and trade when we assume, instead of a zero elasticity of substitution, a unitary elasticity of substitution between inputs in production.¹¹

The change in production that is due to tariffs is most easily seen by relating the change in the price of the primary factor to the expansion of domestic production. Recall that the free trade prices of materials and the output are pegged by the world market prices, and that it is the

elasticity of supply of the primary factor that constrains the expansion of domestic production. In the zero substitution case one way of specifying the expansion of domestic production was by the following (notional) three step procedure: a) determine the effective rate of protection, i.e., the proportionate change in the price of the primary factor (equation (6)); b) determine the proportionate change in the quantity supplied of the primary factor from the elasticity of supply of the primary factor and the effective rate of protection; and c) determine the proportionate change in the quantity of output (equation (9)), given the proportionate change in the quantity supplied of the primary factor and the proportionate change in the number of units of the primary factor per unit of output, the latter being zero in the case of no substitution.

Consider a similar three step procedure in the case of production function homogeneous of degree one with unitary elasticity of substitution, i.e., a Cobb-Douglas production function. First, to calculate the proportionate change in the price of the primary factor, take the Cobb-Douglas total cost function

$$(13) \quad S_j p_j = \frac{\left(\prod_i p_i^{a_{ij}} \right) p_f^{v_j} S_j}{k}$$

where $k = K(a_{ij}^{a_{ij}} \cdot v_j^{v_j})$, which is a constant, and p represents per unit price. Rearranging (13)

$$(14) \quad p_{jk} = \left(\prod_i p_i^{a_{ij}} \right) p_f^{v_j},$$

and since our concern is with relative price changes, set $p_{jk} = 1$, $p_i = 1$ for all i , and $p_f = 1$. Inflating the various prices due to tariffs

$$(15) \quad 1 + t_j = \left[\prod_i (1 + t_i)^{a_{ij}} \right] (1 + f_j)^{v_j}$$

¹⁰ This holds regardless of the degree of substitution in production. Note also that discovery of input tariffs and the effective rate of protection does not mean discovery of any heretofore hidden costs due to tariffs. Rather, we are better able to disentangle the distribution of those cost increases permitted by the nominal tariff. Contrast, for example, Arthur Smith in the introduction of J. R. Melvin and B. W. Wilkinson: "The effective tariff rates derived in this study suggest that cost and productivity effects of the tariff structure in Canada may be significantly larger than those indicated by the nominal rates."

¹¹ The question of substitution was raised initially by Corden (1966) and Travis (1964). Subsequent contributions included: James Anderson and Seija Naya, Corden forthcoming, Grubel and Lloyd, Leith (1968), Lloyd, Ramaswami and Srinivasen, Augustine H. H. Tan, and Travis (1968).

and solve for f_j ¹²

$$(16) \quad f_j = \left[\frac{1 + t_j}{\prod_i (1 + t_i)^{a_{ij}}} \right]^{1/v_j} - 1$$

Second, given f_j from (16), and ϵ_j , dS_{fj}/S_{fj} is obtained. Third, determine $d\alpha_{fj}/\alpha_{fj}$, and substitute both dS_{fj}/S_{fj} and $d\alpha_{fj}/\alpha_{fj}$ into (9) to obtain dS_j/S_j . The expression for $d\alpha_{fj}/\alpha_{fj}$ is obtained by expanding the relationship

$$(17) \quad \alpha_{fj} = v_j \frac{p_j}{p_f}$$

With the introduction of tariffs, (17) becomes

$$(18) \quad \alpha_{fj}(1 + d\alpha_{fj}/\alpha_{fj}) = \frac{v_j p_j (1 + t_j)}{p_f (1 + f_j)}$$

Since (17) holds, cancelling on both sides of (18) yields

$$(19) \quad 1 + d\alpha_{fj}/\alpha_{fj} = \frac{(1 + t_j)}{(1 + f_j)}$$

And hence from (19)

$$(20) \quad d\alpha_{fj}/\alpha_{fj} = \frac{1 + t_j}{1 + f_j} - 1$$

¹² Note that allowing a positive elasticity of substitution ($\sigma > 0$) between inputs results in a higher effective rate of protection than in the zero elasticity of substitution case, where the effective rate of protection is defined as the proportionate change in the price of the primary factor input. (See Leith (1968) p. 600.) If, however, the effective rate of protection is defined as the proportionate change in per unit value-added, no statement can be made about the bias introduced by allowing substitution unless the tariffs and elasticity of substitution are known. And, if the effective rate of protection is defined in this way, it is no longer an indicator of resource pulls. This is a further reason for confining our attention to partial equilibrium analysis. In a general equilibrium model neither the price nor the per unit value-added definition of the effective rate of protection is an indicator of resource pull. The value-added definition in a general equilibrium model suffers from the same problem arising from substitution as in the partial equilibrium model. The price definition has no meaning in a general equilibrium context because domestic factor prices are identical between all activities with and without protection, and hence the proportionate changes in the factor prices are identical between all activities.

Thus, for the Cobb-Douglas case¹³ equation (9) becomes

$$(21) \quad \begin{aligned} dS_j/S_j &= \frac{1 + \epsilon_{fj}}{(1 + t_j)/(1 + f_j)} - 1 \\ &= \frac{(1 + \epsilon_{fj})(1 + f_j)}{1 + t_j} - 1 \end{aligned}$$

When substitution is allowed, economizing of more expensive inputs becomes possible. To illustrate the quantitative significance of substitution in determining the magnitude of the expansion of domestic production, a numerical example is contained in Table 1. Economizing takes place in the unitary elasticity of substitution case ($\sigma = 1$) whenever *relative* prices of inputs change. Thus, in every case, except where all tariffs are the same leaving relative prices of all inputs unchanged (col. 5), the effective rate of protection and the change in production are different between $\sigma = 1$ and $\sigma = 0$. Further, note that when the tariff rate on the output (t_j) exceeds the average tariff rate on the inputs (\bar{t}_i , where $\bar{t}_i = \sum_i a_{ij} t_i / a_{ij}$), there is economizing of the primary factor, and where $t_j < \bar{t}_i$, relatively more of the primary factor is used.¹⁴ As a result of the economizing, the absolute rate of change in domestic output is greater where economizing takes place than when it is not permitted. Note also that the combination of tariffs that yields a zero effective rate of protection is different between the two cases: for $\sigma = 0$, $f_j = 0$ when $t_j = \sum_i a_{ij} t_i$ (col. 4); and for $\sigma = 1$, $f_j = 0$ when

$$(1 + t_j) = \prod_i [(1 + t_i)^{a_{ij}}]$$

¹³ The same approach can be followed using the effective rate of protection calculated from the CES production function. (See Leith 1968, Part iii.) The procedure is: (a) calculate the effective rate of protection using equation (12) in *ibid.*; and (b) determine $d\alpha_{ij}/\alpha_{ij}$ working from the input demand equation (10) in *ibid.*; and (c) calculate dS_j/S_j from equation (9) of this paper.

¹⁴ For a geometric representation of substitution between the primary factor and material inputs, see Leith (1967.)

TABLE 1—NUMERICAL EXAMPLE OF ERP AND PRODUCTION CHANGE UNDER ZERO AND UNITARY ELASTICITY OF SUBSTITUTION ASSUMPTIONS

(Free trade coefficients: $a_{1j} = .33$, $a_{2j} = .33$, $v_j = .33$, and assumed $\epsilon_j = .33$)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
t_j	output	.20	.20	.20	.20	.20	.20	.20	.20	.20	.20
t_1	input 1	.50	.40	.50	.30	.20	.10	.10	.05	0	-.10
t_2	input 2	.50	.40	.30	.30	.20	.10	0	.05	0	+.10
\bar{t}_i	av. input tariff	.50	.40	.40	.30	.20	.10	.05	.05	0	0
f_j	$\sigma=0$	-.400	-.20	-.20	0	.20	.40	.50	.50	.60	.60
f_j	$\sigma=1$	-.232	-.118	-.114	+.0225	.20	.428	.571	.567	.728	.745
$d\alpha_{ij}/\alpha_{ij}$	$\sigma=0$	0	0	0	0	0	0	0	0	0	0
$d\alpha_{ij}/\alpha_{ij}$	$\sigma=1$	+.563	+.361	+.354	+.174	0	-.160	-.236	-.234	-.306	-.312
dS_j/S_j	$\sigma=0$	-.133	-.067	-.067	0	+.067	+.133	+.167	+.167	+.20	+.20
dS_j/S_j	$\sigma=1$	-.410	-.294	-.290	-.141	+.067	+.361	+.558	+.552	+.791	+.814

Note: \bar{t}_i is the weighted average input tariff: i.e., $\bar{t}_i = \sum_j a_{ij}t_i / \sum_j a_{ij}$

Where there are input tariffs two effects must be distinguished: substitution between material and primary inputs; and substitution between material inputs. If all material inputs are subject to the same tariff, the substitution is only between the factor and the materials, and not between materials (e.g., col. 8). However, where inputs are subject to different tariffs, there is economizing of the relatively more expensive inputs, and the cost-increasing effect of input tariffs is not as strong, the effective rate of protection is (algebraically) greater, and the change in production is (algebraically) greater (e.g., col. 7 vs. 8, and col. 3 vs. 2). Further, an import subsidy on one input to offset an import tariff on another is *not* equivalent to identical tariffs on both inputs yielding the same average input tariffs. For example, $t_1 = t_2 = 0$ (col. 9) is not equivalent to $t_1 = -.10$ and $t_2 = .10$ so that $\bar{t}_i = 0$ (col. 10), for substitution in the latter case between inputs 1 and 2 means that the effective rate of protection is greater and the expansion of production is greater than in the former case.

In the expansion of J production we draw not only primary factors, but also material inputs. In a manner similar to the derivation of (21) we can solve for $d\alpha_{ij}/\alpha_{ij}$ from the relationship $\alpha_{ij} = a_{ij}p_j/p_i$ with the result

$$(22) \quad \frac{d\alpha_{ij}}{\alpha_{ij}} = \frac{1 + t_j}{1 + t_i} - 1$$

Then, from $S_{ij} = S_j\alpha_{ij}$ we can solve for

$$(23) \quad \frac{dS_{ij}}{S_{ij}} = \frac{(1 + \epsilon_{ij})(1 + f_j)}{1 + t_i} - 1$$

In addition to the change in production, consider now the change in consumption and the change in use of good J in the case of unitary rather than zero elasticity of substitution.

Clearly the change in consumption is unaffected by the elasticity of substitution in production. Thus, equation (3) still holds.

The change in use equation, however, is altered by the introduction of substitution, because the drawing of J into the various other I activities is affected by the elas-

ticity of substitution in the latter's production functions. Thus, in a manner similar to (23), and with some rearranging, the change in use of good J is

$$(24) \quad dU_j = \sum_i a_{ji} \left[\frac{(1 + \epsilon_{ji})(1 + f_i)}{1 + t_j} - 1 \right] S_i$$

Finally, summing all the components of the change in imports, when $\sigma=1$,

$$(25) \quad dM_j = - \left[\frac{(1 + \epsilon_{ji})(1 + f_i)}{1 + t_j} - 1 \right] S_i \\ + \eta_j t_j C_j \\ + \sum_i a_{ji} \left[\frac{(1 + \epsilon_{ji})(1 + f_i)}{1 + t_j} - 1 \right] S_i$$

Comparing the change in imports due to tariffs in the case of $\sigma=1$ with the case of $\sigma=0$ shows that: a) the component measuring the change in domestic output of J will be greater for $\sigma=1$ when $f_i > t_j$; b) the component measuring the change in consumption will be the same between $\sigma=1$ and $\sigma=0$; and c) the component measuring the change in use of J by the typical using industry I will be greater for $\sigma=1$ when $f_i > t_j$.

REFERENCES

- J. Anderson and S. Naya, "Substitution and Two Concepts of Effective Rate of Protection," *Amer. Econ. Rev.*, Sept. 1969, 59, 607-12.
- B. Balassa, "The Impact of the Industrial Countries' Tariff Structure on Their Imports of Manufactures from Less Developed Areas," *Economica*, Nov. 1967, 34, 372-83.
- , "Tariff Protection in Industrial Countries: An Evaluation," *J. Polit. Econ.*, Dec. 1965, 73, 572-94.
- W. M. Corden, "The Structure of a Tariff System and the Effective Protective Rate," *J. Polit. Econ.*, June 1966, 74, 221-37.
- , "The Substitution Problem in the Theory of Effective Protection," in *Theory of Protection*, forthcoming.
- R. Dardis, "Intermediate Goods and the Gain from Trade," *Rev. Econ. Statist.*, Nov. 1967, 49, 502-09.
- H. G. Grubel and P. J. Lloyd, "Factor Substitution and Effective Tariff Rates," *Rev. Econ. Stud.*, forthcoming.
- H. G. Johnson, "The Theory of Tariff Structure with Special Reference to World Trade and Development," in *Trade and Development*, Geneva 1965.
- , "The Theory of Effective Protection and Preferences," *Economica*, May 1969, 36, 119-38.
- J. C. Leith, "Substitution and Supply Elasticities in Calculating the Effective Protective Rate," *Quart. J. Econ.*, Nov. 1968, 82, 588-601.
- , "Effective Rates of Protection: Analysis and an Empirical Test," unpublished doctoral dissertation, Univ. Wis., 1967.
- and G. L. Reuber, "The Impact of the Industrial Countries' Tariff Structure on their Imports of Manufacturers from Less Developed Areas: A Comment," *Economica*, Feb. 1969, 36, 75-80.
- P. J. Lloyd, "Effective Protection: A Different View," *Econ. Rec.*, forthcoming.
- R. I. McKinnon, "Intermediate Products, and Differential Tariffs: A Generalization of Lerner's Symmetry Theorem," *Quart. J. Econ.*, Nov. 1966, 80, 584-615.
- B. F. Massell, "The Resource-Allocative Effects of a Tariff and the Effective Protection of Individual Inputs," *Econ. Rec.*, Sept. 1968, 44, 369-76.
- J. R. Melvin and B. W. Wilkinson, *Effective Protection in the Canadian Economy*, Economic Council of Canada Special Study No. 9, Ottawa 1968.
- V. K. Ramaswami and T. N. Srinivasan, "Tariff Structure and Resource Allocation in the Presence of Factor Substitution," in J. N. Bhagwati et al, eds., *Trade, Balance of Payments, and Growth; Essays in Honor of Charles Kindleberger*, Cambridge, Mass. 1971.
- R. J. Ruffin, "Tariffs, Intermediate Goods, and Domestic Protection," *Amer. Econ. Rev.*, June 1969, 59, 261-69.
- A. H. H. Tan, "Differential Tariffs, Negative Value-Added and the Theory of Effective

- Protection," *Amer. Econ. Rev.*, Mar. 1970, 60, 107-16.
- W. P. Travis, *The Theory of Trade and Protection*, Cambridge, Mass. 1964.
- , "The Effective Rate of Protection and the Question of Labor Protection in the United States," *J. Polit. Econ.*, May-June 1968, 76, 443-61.
- P. Wonnacott and R. J. Wonnacott, *Free Trade Between the United States and Canada: The Potential Economic Effects*, Cambridge, Mass. 1967.

A General Disequilibrium Model of Income and Employment

By ROBERT J. BARRO AND HERSCHEL I. GROSSMAN*

As is now well understood, the key to the Keynesian theory of income determination is the assumption that the vector of prices, wages, and interest rates does not move instantaneously from one full employment equilibrium position to another. By implication, Keynesian economics rejects the market equilibrium framework for analyzing the determination of quantities bought, sold, and produced. This framework is associated with Walras and Marshall, both of whom proceeded as if all markets were continuously cleared. Walras rationalized this procedure by incorporating recontracting arrangements, while Marshall did so by regarding price adjustments to be an instantaneous response to momentary discrepancies between quantities supplied and demanded.

By rejecting these rationalizations, Keynesian theory proposes as a general case a system of markets which are not always cleared. Keynes was, tacitly at least, concerned with the general theoretical problem of the intermarket relationships in such a system. The failure of a market to clear implies that, for at least some individuals, actual quantities transacted diverge from the quantities which they supply or demand. Thus, the natural focus of Keynesian analysis is on the implications for behavior in one market of the existence of such a divergence in another market. Indeed, some recent writers, such as Robert Clower and Axel Leijonhufvud, have argued very convincingly that this

focus is the crucial distinguishing feature of Keynesian economics.

Unfortunately, the evolution of conventional post-Keynesian macroeconomics failed to interpret the Keynesian system in this light.¹ Instead, conventional analysis has chronically attempted to coax Keynesian results out of a framework of general market equilibrium. The result has been to leave conventional macroeconomics with an embarrassingly weak choice-theoretic basis, and to associate with it important implications which are difficult to reconcile with observed phenomena.

A classic example of such a difficulty concerns the relationship between the level of employment and the real wage rate. In the conventional analysis, the demand for labor is inversely and uniquely related to the level of real wages. This assumption accords with Keynes; who, in this respect, had adhered to received pre-Keynesian doctrine.² Given this assumption, cyclical variations in the quantity of labor demanded and the amount of employment must imply countercyclical variation in real wage rates. As is well known, however, such a pattern of real wages has not been observed.³

¹ See Leijonhufvud.

² Keynes wrote:

... with a given organization, equipment and technique, real wages and the volume of output (and hence of employment) are uniquely correlated, so that, in general, an increase in employment can only occur to the accompaniment of a decline in the rate of real wages. Thus, I am not disputing this vital fact which the classical economists have (rightly) asserted. ... The real wage earned by a unit of labor has a unique inverse correlation with the volume of employment. [1936, p. 17]

³ The evidence has been recently reviewed by Edwin

* Assistant professor and associate professor of economics, respectively, Brown University. National Science Foundation Grants GS-2419 and GS-3246 supported this research.

A few authors have pointed out the inappropriateness of attempts to force Keynesian analysis into a market equilibrium framework. Contributions by Don Patinkin (1956) and Clower, in particular, represent important attempts to reconstruct macroeconomic theory within an explicitly disequilibrium context.

In the unfortunately neglected chapter 13 of *Money, Interest, and Prices*, Patinkin analyzed involuntary unemployment in a context of explicit market disequilibrium; and he showed that the misleading implications of the conventional analysis regarding the real wage are a direct consequence of its general equilibrium character.⁴ Patinkin presented a theory in which involuntary unemployment of labor can arise as a consequence of disequilibrium, in particular, excess supply in the market for current output. In this theory, the inability of firms to sell the quantity of output given by their supply schedule causes them to demand a smaller quantity of labor than that given by their conventional (or notional) demand schedule. The immediate significance of this theory is that it is able to generate unemployment without placing any restrictions on the level or movement of the real wage.⁵ Unemploy-

ment of labor requires only that the vector of prices and wages implies a deficiency of demand for current output. As Patinkin suggests, this interpretation of the proximate cause of unemployment is more Keynesian than Keynes' own discussion.

The essence of Patinkin's theory is causality running from the level of excess supply in the market for current output to the state of excess supply in the market for labor. Patinkin thereby explains the proximate cause of cyclical unemployment, but his analysis involves only partial, rather than general, disequilibrium. At the least, a general disequilibrium model would, in addition, incorporate the possibility of a reverse influence of the level of excess supply in the labor market upon the state of excess supply in the market for current output.

Clower's important paper develops a theory emphasizing this causal relationship. He presents a derivation of the Keynesian consumption function in which he interprets the relationship between consumption and income as a manifestation of disequilibrium in the labor market. This approach to explaining household behavior is obviously similar to Patinkin's analysis of the firm. The only significant difference is that Clower's households have a choice between consuming and saving, so that his problem is explicitly choice theoretic. However, if Patinkin's approach were generalized to a multi-input production function, the resulting analysis would be formally analogous to Clower's.

The analysis in this paper builds on the foundations laid down by the Patinkin and Clower analyses of a depressed economy. Our purpose is to develop a generalized analysis of both booms and depressions as disequilibrium phenomena.⁶ Section I

Kuh, esp. pp. 246-48; and Ronald Bodkin. Keynes (1939) recognized this discrepancy, and offered a rather contrived explanation for it in terms of monopoly and procyclical variation in demand elasticities. More recently, Kuh attempted to explain this discrepancy in terms of a fixed proportions production function in the short run.

⁴ Chapter 13 also appears, apparently unchanged, in the second edition of *Money, Interest, and Prices* (1965). Patinkin had first presented some of the essentials of this analysis in an earlier article (1949). A similar formulation appears in Edgar Edwards.

⁵ Patinkin's theory does not involve any restrictions either upon the substitutability among factors of production or upon demand elasticities. (See fn. 3.) Of course, this theory does not deny that an excessive level of real wages can be an independent cause of unemployment. But, a clear analytical distinction is made between unemployment due to this cause, and unemployment which occurs even when the level of real wages is not excessive.

⁶ The analysis by Robert Solow and Joseph Stiglitz, although they emphasize different questions, is somewhat similar to the present approach. However, their analytical format does differ from ours in at least three

sketches the analytical framework employed. Section II reviews and generalizes Patinkin's analysis of the labor market and involuntary unemployment. Section III develops a distinction, implied by Patinkin's analysis, between two concepts of unemployment; one associated with excess supply in the labor market and the other associated with equilibrium in the labor market but with disequilibrium elsewhere in the system. Section IV reviews Clower's analysis and shows how it is formally analogous to Patinkin's. Section V joins the Patinkin and Clower analyses into a model of an economy experiencing deficient aggregate demand. Section VI formulates an analogous model of an economy experiencing excessive aggregate demand. Finally, Section VII summarizes the main results.

I. Analytical Framework

The following discussion utilizes a simple aggregative framework which involves three economic goods—labor services, consumable commodities, and fiat money—and two forms of economic decision making unit—firms and households. Labor services are the only variable input into the production process. Other inputs have a fixed quantity, no alternative use, and zero user cost. Consumable commodities are the only form of current output; there is no investment.⁷ Money is the only store of value, and it also serves as a medium of exchange and unit of account. The nominal quantity of money is exogenous and constant.

Firms demand labor and supply commodities. They attempt to maximize

substantial respects: First, they do not discuss the choice-theoretic basis for the theory. Second, the equilibrium price level is indeterminate in their model. Third, by introducing restrictions on the rate of change of employment, they complicate matters and obscure what would seem to be essential in the intermarket effects of disequilibrium.

⁷ It should be clear that the incorporation of investment and a market for securities would alter none of the conclusions advanced in this paper.

profits. Households supply labor and demand commodities and money balances. They also receive the profits of the firms according to a predetermined distribution pattern. Households attempt to maximize utility. Each firm and household is an atomistic competitor in the markets for both commodities and labor.

Following Patinkin (1956, 1965), each of the flow variables in the model—commodities, labor services, and the increment to money balances—is for simplicity expressed as the quantity which accrues over a finite unit of time, say a week, so that each assumes the dimensions of a stock. The model thus includes the following variables:

- y = quantity of commodities
- x = quantity of labor services
- m = increment to real money balances (in commodity units)
- π = quantity of real profits (in commodity units)
- M = initial stock of nominal money balances
- P = money price of commodities
- w = real wage rate (in commodity units)

Throughout the following discussion, the method of analysis is to take a particular vector of the price level and real wage rate as given, and to work out the levels of income and employment implied by that vector. This procedure represents a non-Marshallian, or Keynesian, extreme, and following John Hicks may be denoted as the "fix-price method." The analysis does, of course, have implications for the appropriate specification of the forces making for changes in prices and wages. This paper does not explicitly investigate these implications, although we do consider a parenthetical example concerning the model's implications for the cyclical behavior of real wages.⁸

⁸ Grossman develops a more general model of multi-market disequilibrium based on Clower's choice-theoretic paradigm, and focuses in detail on the implications

II. Patinkin's Analysis of the Labor Market

Consider the behavior of the representative firm under the provisional assumption that it regards profit maximization as being constrained only by the production function. In particular, the firm perceives that it can purchase all the labor which it demands and sell all the output which it supplies at the existing levels of w and P . Thus, profits are given by

$$\pi = y^S - wx^D,$$

where the superscripts indicate supply and demand quantities. Assuming the production function to be

$$y = F(x),$$

with positive and diminishing marginal product, profit maximization implies

$$x^D = x^D(w),$$

such that $\partial F/\partial x = w$, and

$$y^S = F(x^D)$$

Patinkin (1956, 1965) contrasts the above to a situation in which commodities are in excess supply. Voluntary exchange implies that actual total sales will equal the total quantity demanded. The representative firm will not be able to sell its notional supply y^S .⁹ Let y represent its actual demand-determined sales, where $y < y^S$.¹⁰ Then, the profit maximization

of this model for the disequilibrium behavior of prices and interest.

⁹ We assume here that the firm would actually like to sell y^S . Such behavior may not always be optimal. For example, Section VI discusses a situation of excess demand for labor in which the firm's effective supply $y^{S'}$ is less than y^S . However, we assume for simplicity that excess demand for labor never coexists with excess supply of commodities and vice versa. Grossman presents a more general treatment of multi-market disequilibria which allows for the coexistence of excess supply in one market and excess demand in another, as well as excess supply or demand in both.

¹⁰ In principle, y need not be less than y^S for every firm. The apportionment of the actual sales among the firms depends upon established queuing or rationing procedures. Grossman presents an explicit analysis of this apportionment within a framework of voluntary exchange.

problem becomes simply to select the minimum quantity of labor necessary to produce output quantity y .¹¹ In other words, the firm maximizes

$$\pi = y - wx^{D'},$$

subject to $y = F(x)$. The variable $x^{D'}$ may be denoted as the effective demand for labor. Profit maximization now implies

$$(1) \quad x^{D'} = F^{-1}(y) \quad \text{for} \quad \frac{dF}{dx} \geq w$$

The constraint of $y < y^S$ implies $x^{D'} < x^D$, with $x^{D'}$ approaching x^D as y approaches y^S .¹²

The inability of a firm to sell its desired output at the going price violates an assumption of the perfectly competitive model. Kenneth Arrow has stressed this inconsistency of perfect competition with disequilibrium. Essentially, he argues that economic units which act as perfect competitors in equilibrium must (at least in certain respects) perform as monopolists in disequilibrium. In this paper we focus on the reaction of economic units to given (equilibrium or disequilibrium) price levels. If, in addition, one wished to analyze explicitly the dynamics of price adjustment, it would be necessary to discard the perfectly competitive paradigm of the producer as a price taker. (In this regard, see Barro 1970, 1971.)

¹¹ This analysis abstracts from inventory accumulation or decumulation. For simplicity, we assume throughout that output always adjusts instantaneously to equal the smaller of supply and demand. Permitting inventory accumulation would not affect the essentials of the analysis, although it would introduce a complication analogous to the inclusion of an additional input. In general, we might obtain $dy/dt = k[\min(y^D, y^S) - y]$, where $k = k(w, y) > 0$. A similar gradual adjustment process for employment might also be possible, as in Solow and Stiglitz.

¹² The choice-theoretic nature of the problem becomes much more interesting when there is more than one form of input. Assume profits to be given by $\pi = y - w_1x_1^{D'} - w_2x_2^{D'}$, where the production function is $y = F(x_1, x_2)$, which has the usual convexity properties. Profit maximization now implies

$$(1.1) \quad x_1^{D'} = x_1^{D'}\left(\frac{w_1}{w_2}, y\right)$$

$$(1.2) \quad x_2^{D'} = x_2^{D'}\left(\frac{w_1}{w_2}, y\right)$$

such that at output y , $(\partial F/\partial X_1)/(\partial F/\partial X_2) = (w_1/w_2)$. In reducing output y^S to y , the firm must now make a decision regarding optimal input combinations. However, as y approaches y^S , $x_1^{D'}$ and $x_2^{D'}$ approach x_1^D and x_2^D .

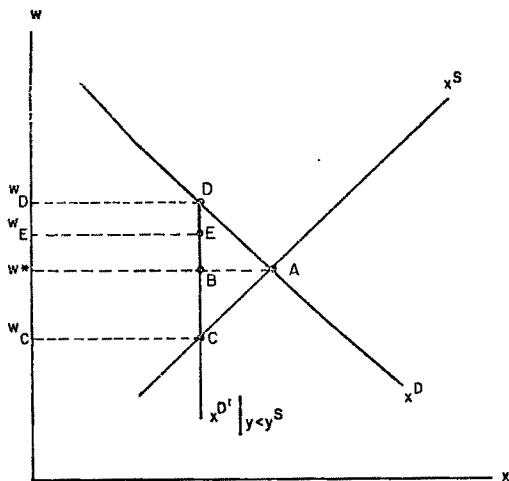


FIGURE 1. THE LABOR MARKET WITH EXCESS SUPPLY OF COMMODITIES

The essential implication of equation (1) is that the effective demand for labor can vary even with the real wage fixed. Given voluntary exchange, employment cannot exceed the effective demand for labor. The quantity of employment thus is not uniquely associated with the real wage.

III. The Concept of Unemployment

Figure 1 depicts the preceding analysis of the labor market. The notional demand schedule for labor x^D is downward sloping. If $y = y^S$, the effective demand for labor $x^{D'}$ coincides with the notional demand. If $y < y^S$, the effective demand is independent of the real wage and less than the notional demand. The (notional) supply schedule for labor x^S , which will be derived below, is shown as upward sloping.

Figure 1 suggests a distinction between two concepts of unemployment—involuntary unemployment associated with excess (effective) labor supply, and voluntary unemployment associated with equilibrium in the labor market, but with disequilibrium elsewhere in the system. Suppose that initially the commodity market is in equilibrium, so that $y = y^S$ and $x^{D'} = x^D$, and that initially the real wage is w^* . Thus, the labor market is in equilib-

rium at point A, which may be denoted as full employment general equilibrium. Now suppose, say because the price level P is too high, that commodity demand is lower so that $y < y^S$ and $x^{D'} < x^D$. At the real wage w^* , excess supply of labor will amount to quantity AB . Failure of the price level to adjust to clear the commodity market leads to excess supply in the labor market. This excess supply represents what we usually refer to as involuntary unemployment. It is also what the Bureau of Labor Statistics ideally intends to represent by its statistical measure of unemployment—those seeking but not obtaining work at the going real wage. Involuntary unemployment clearly does not require a rise in the real wage above the level consistent with full employment equilibrium.

Now suppose that the real wage were to decline to w_C , so that the supply and effective demand for labor are equilibrated at point C. At point C, involuntary unemployment has vanished, but clearly this situation is not optimal. The reduced real wage has induced AB man-hours of labor to leave the labor force. Employment remains AB man-hours below the level associated with general equilibrium. Involuntary, i.e., excess supply, unemployment has been replaced by voluntary unemployment.¹³

The conclusion is that too high a real wage was not the cause of the lower employment, and a reduction in the real wage

¹³ In terms of the *BLS* unemployment statistic, it is not clear that "zero" unemployment would be measured at w_C . If the higher wage, w^* , were (at least for a time) viewed as "normal," a considerable proportion of job seekers at wage w_C would be those willing to work at w^* , but not at w_C . These people are in the labor market seeking information on possible employment opportunities at (or above) w^* , and would not actually be willing to work at the going wage (see Armen Alchian). To the extent that the *BLS* measure includes this type of frustrated job seeker, the index will be a better measure of the gap between actual and general equilibrium employment BA , while simultaneously being a poorer index of those seeking but not obtaining employment at the going wage w_C .

is only a superficial cure. The real cause of the problem was the fall in commodity demand, and only a reflation of commodity demand can restore employment to the proper level.

The above analysis suggests the following cyclical patterns of real wages and employment: A decline in commodity demand and output produces a decline in employment with a corresponding excess supply of labor (point *B*). To the extent that real wages decline in response to this excess supply, a fall in real wages toward w_c will accompany (follow upon) the decline in employment. If, at point *C* or at some intermediate point between *B* and *C*, some action is taken to restore effective commodity demand, excess demand for labor (or, at least reduced excess supply) will result. In that case, a rising real wage may accompany the recovery of output and employment. Thus, disequilibrium analysis of the labor market suggests that real wages may move procyclically. This result differs from the conventional view that employment and real wages must be inversely related.

The present model can also be used to analyze involuntary unemployment which results from an excessive real wage. Clearly, if the real wage were above w^* , no stimulation of commodity demand could bring about full employment equilibrium, unless the real wage were reduced. This classical type of involuntary unemployment should be clearly distinguished from the type of unemployment discussed above, which arises, with the real wage at or below w^* , from a deficiency of demand for commodities.

IV. Clower's Analysis of the Consumption Function

In order to close the model, we must also analyze household behavior. Consider the behavior of the representative household under the provisional assumption that it regards utility maximization as being sub-

ject only to the budget constraint. In particular, the household perceives that it can sell all the labor which it supplies and purchase all the commodities which it demands at the existing levels of w and P . Assume the utility function to be

$$U = U\left(x^s, y^D, \frac{M}{P} + m^D\right),$$

with the partial derivatives $U_1 < 0$, $U_2 > 0$, and $U_3 > 0$. The budget constraint is

$$\pi + wx^s = y^D + m^D$$

x^s , y^D , and m^D may be denoted as the notional supply of labor, the notional demand for commodities, and the notional demand for additional money balances. Utility maximization in general will imply that x^s , y^D , and m^D are each functions of w , M/P , and π . For simplicity, we shall assume that x^s depends only on the real wage. The important point is that the notional demand functions for commodities and additional money balances do not have the forms of the usual consumption and saving functions with income as an argument, because the household simultaneously chooses the quantity of labor to sell.

Clower contrasts the above notional process to a situation in which labor services are in excess supply. Given voluntary exchange, actual total employment in this situation equals the total quantity demanded. Thus, the representative household is unable to sell its notional labor supply x^s and obtain its implied notional labor income wx^s .¹⁴ Labor income is no longer a choice variable which is maximized out, but is instead exogenously given. We may assume that the representative household is able to obtain the quantity of employment x , where $x < x^s$,

¹⁴ We assume that the household would actually like to sell x^s . As indicated in fn. 9, we assume for simplicity that excess demand for commodities never coexists with excess supply of labor.

so that its total income is $w\pi + \pi$. In this case, the utility maximization problem amounts to the optimal disposition of this income.

In other words, the household maximizes

$$U\left(x, y^{D'}, \frac{M}{P} + m^{D'}\right)$$

subject to $\pi + wx = y^{D'} + m^{D'}$. The variables $y^{D'}$ and $m^{D'}$ may be denoted as the effective demands for commodities and additional money balances. Utility maximization now implies

$$(2) \quad y^{D'} = y^D \left(\pi + wx, \frac{M}{P} \right),$$

and

$$(3) \quad m^{D'} = m^D \left(\pi + wx, \frac{M}{P} \right)$$

Note that, in aggregate, $\pi + wx = y = F(x)$. Thus, since all income accrues to the households, consumption and saving demand depend ultimately only on the level of employment and real money balances and not on the real wage rate. The constraint $x < x^S$ would generally imply $y^{D'} < y^D$ and $m^{D'} < m^D$, but as x approaches x^S , $y^{D'}$ and $m^{D'}$ approach y^D and m^D .¹⁵

The important property of equations (2) and (3) is that they do have the form of the usual Keynesian consumption and saving functions. Labor income enters the consumption and saving functions as it represents the constraint upon the demand for current output imposed by the excess supply of labor.

¹⁵ To the extent that long-run employment (income) exceeds current employment (income), a household may be more willing to maintain a higher demand for commodities at the expense of money balances. In this case effective commodity demand would remain closer to notional demand, and the "income multiplier" (as depicted later in Figure 4) would be smaller. In general, the size of the effect of quantity constraints on effective demands will depend on whether the constraint is viewed as "permanent" or "transitory."

The formal analogy between the Clower and Patinkin models should be apparent from the derivations of equations (2), (3) and equation (1), or more particularly equations (1.1) and (1.2) in footnote 12. Patinkin's model involves profit maximization subject to an output constraint, whereas Clower's model involves utility maximization subject to an employment constraint.

V. General Disequilibrium Involving Excess Supply

In Patinkin's analysis, the effective demand for labor was derived for a given level of demand for current output. To close this model, the demand for current output must be explained. In Clower's analysis, the effective demand for current output was derived for a given level of demand for labor. To close this model, the demand for labor must be explained. Thus, the Patinkin and Clower analyses are essential complements. When appropriately joined, they form a complete picture of the determination of output and employment in a depressed economy.

Figure 2 depicts Clower's analysis of the commodity market. The notional supply schedule for commodities is a downward sloping function of the real wage. The two notional demand schedules are upward sloping functions, reflecting the effect of substitutability between consumption and leisure as well as a positive income effect. As the real wage rate rises, leisure becomes relatively more expensive, and households tend to work and consume more. The schedule corresponding to the general equilibrium price level P^* passes through the point A . At point A , which corresponds to point A in Figure 1, P^* and w^* are consistent with simultaneous notional equilibrium in both the labor and commodity markets. The other notional commodity demand schedule in Figure 2 corresponds to the higher price level P_1 . Because of the real balance effect, this

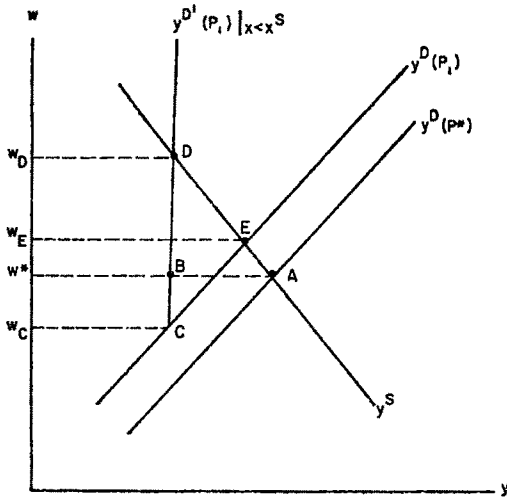


FIGURE 2. THE COMMODITY MARKET WITH EXCESS SUPPLY OF LABOR

curve lies to the left of the curve associated with P^* .¹⁶ If $x = x^S$, the effective demand for commodities coincides with the notional demand. If $x < x^S$, the effective demand is independent of the real wage, as noted above, and is less than the notional demand. The effective demand schedule shown in Figure 2 corresponds to the higher price level P_1 . Points B , C , D , and E also correspond to the same points in Figure 1. This correspondence can be seen most clearly by explicitly depicting the interaction between the two markets, as is done in Figures 3 and 4.

Figure 3 illustrates the relationship between the existence of excess supply in one market and the other. In Figure 3, the points A , B , C , D , and E coincide with the same points in Figures 1 and 2. The four loci separate the regions of inequality between the indicated supply and demand

¹⁶ As the model is constructed, only y^D and m^D of the five notional schedules; x^D , x^S , y^S , and m^S depend on the price level independently of the real wage. In a more general model, real balances would affect x^D , x^S , and y^S , and the price level would affect these schedules also. By ignoring this possibility, the exposition is simplified without losing any of the essence of the analysis. Of course, if none of the five schedules were influenced by the price level, prices would not be determined within the model.

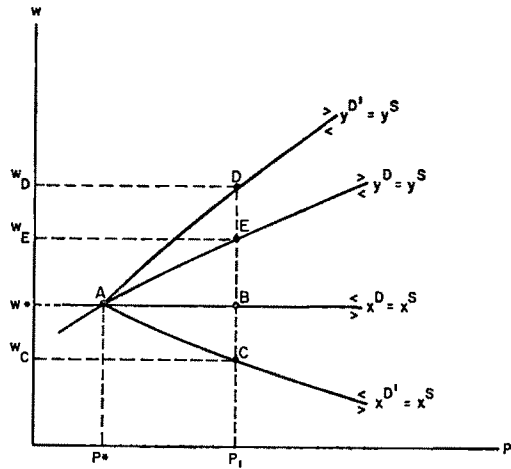


FIGURE 3. INTERACTION OF EXCESS SUPPLY IN BOTH MARKETS

concepts. The locus $x^D = x^S$ is horizontal because, by assumption, both x^D and x^S depend only on the real wage. The locus $y^D = y^S$ is upward sloping because as shown in Figure 2, y^S is a decreasing function of the real wage, whereas y^D is an increasing function of the real wage (substitution and income effect) and a decreasing function of the price level (real balance effect). These loci intersect at point A , which depicts full employment general equilibrium. Points B , C , D , and E are all associated with a price level P_1 , which is higher than the equilibrium price level P^* .¹⁷ Point B , for example, would be consistent with notional equilibrium in the labor market, but implies excess supply in the commodity market. The essential point of Patinkin's analysis is that the effective demand for labor is smaller than the notional demand when commodities are in excess supply. Thus, the locus $x^{D'} = x^S$ exists to the right of point A and lies everywhere below the locus $x^D = x^S$. The existence of excess supply in the commodity market enlarges the region of excess supply in the labor market. Similarly, according to Clower's

¹⁷ We could, of course, just as well think of these points as being associated with a nominal money supply which is too small.

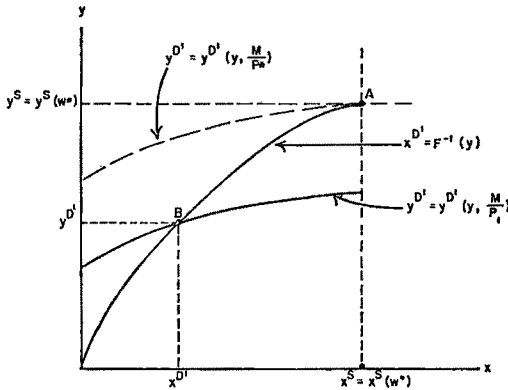


FIGURE 4. OUTPUT AND EMPLOYMENT WITH EXCESS SUPPLY IN BOTH MARKETS

analysis, the effective demand for commodities is less than the notional demand when labor is in excess supply. Thus, the locus $y^{D'} = y^S$ exists to the right of point A and lies everywhere above the locus $y^D = y^S$. The existence of excess supply in the labor market also enlarges the region of excess supply in the commodity market.

Figure 4 illustrates the determination of the actual quantities of current output and employment when there is excess supply in both markets. In particular, Figure 4 has been drawn under the assumption that the existing wage-price vector is (w^*, P_1) , that is that the economy is at point B of Figures 1, 2, and 3. Given voluntary exchange, x and y are determined by $x = \min[x^{D'}, x^S]$ and $y = \min[y^{D'}, y^S]$. The solid locus $x^{D'} = F^{-1}(y)$ describes firm behavior for values of y less than y^S . The solid locus $y^{D'} = y^{D'}(y, M/P_1)$ describes household behavior for values of x less than x^S . The intersection of these two loci determines the values of x and y corresponding to point B . Point A , full employment equilibrium, is at the intersection of y^S and x^S . Since at point B the real wage is consistent with full employment equilibrium, a movement from B to A involves on net only a fall in the price level from P_1 to P^* . In Figure 4, this fall in P is represented by an upward shift in $y^{D'}$ to the dashed locus $y^{D'} = y^{D'}(y, M/P^*)$, which intersects $x^{D'}$ at

point A . The income multiplier in this case is given by the ratio of the difference between y^S and $y^{D'}(B)$ to the vertical distance between the two curves $y^{D'}(P^*)$ and $y^{D'}(P_1)$. Figure 4 is simply the Keynesian cross diagram with employment replacing income on the horizontal axis.

VI. General Disequilibrium Involving Excess Demand

The preceding discussion has concentrated on the case of excess supply in the markets for both commodities and labor. However, analogous considerations clearly apply to the boom situation of excess demand for both commodities and labor.

First, consider the behavior of the representative firm when there is excess demand for labor. The representative firm will be able to obtain the quantity of labor x , where $x < x^D$. The firm then must maximize

$$\pi = y^{S'} - wx$$

subject to $y = F(x)$. The variable $y^{S'}$ may be denoted as the effective supply of commodities. The problem is simply to produce as much output as possible with the available labor. The solution is

$$(4) \quad y^{S'} = F(x) \quad \text{for} \quad \frac{dF}{dx} \geq w$$

Figure 5 depicts the commodity market in this situation, and is analogous to Figure 2. The price level P_2 is assumed to be below P^* .

Next, consider the behavior of the representative household when there is excess demand for commodities. The representative household will be able to obtain the quantity of commodities y , where $y < y^D$. The household then has to choose between either saving, i.e., accumulating as money balances the income which it cannot spend on consumption, or substituting leisure for the unobtainable commodities by supplying less labor, or some combination of the

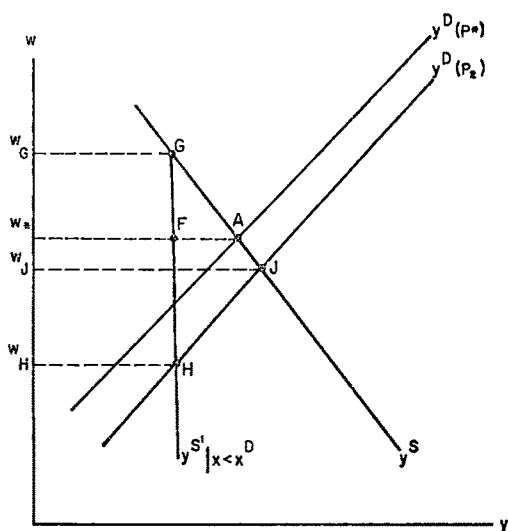


FIGURE 5. THE COMMODITY MARKET WITH EXCESS DEMAND FOR LABOR

two. Formally, the household's problem is to maximize

$$U\left(x^{S'}, y, \frac{M}{P} + m^{D'}\right)$$

$$\text{subject to } \pi + wx^{S'} = y + m^{D'}$$

The variable $x^{S'}$ may be denoted as the effective supply of labor. Utility maximization now implies

$$(5) \quad x^{S'} = x^S\left(w, \frac{M}{P}, \pi, y\right),$$

and

$$(6) \quad m^{D'} = m^{D'}\left(w, \frac{M}{P}, \pi, y\right)$$

This theory stresses the fact that a household may react to frustrated commodity demand in two ways. First, the household may save the income which cannot be spent on consumption (in this model, solely by augmenting money balances). This option corresponds to the classical concept of forced saving, or, more precisely, what D. H. Robertson defined as "automatic lacking." Second, the household may increase leisure by reducing its supply of labor. The second option prob-

ably becomes more important when excess commodity demand is chronic, as in wartime or during other periods of rationing and price controls.¹⁸ However, given that consumption, saving, and leisure in aggregate are substitutes, in general some combination of the two options will always be optimal. Excess demand will generally result in some fall in output.

Classical analysis, in which labor supply is solely a function of the real wage, assumes that households channel all frustrated commodity demand into forced saving. The possibility of reduced labor supply is ignored. However, the inclusion of this option is especially interesting, since it has the apparently paradoxical implication that excess commodity demand can result in decreased employment and output.

Figure 6, which is analogous to Figure 1, depicts the labor market in this situation. Two important observations should be stressed. First, too low a real wage, that is a real wage below the level consistent with general equilibrium, is not a necessary condition for excess demand for labor, even though the notional demand and supply for labor are both assumed to depend only upon the real wage. This observation is obviously the converse of the earlier observation that the effective demand for labor is not uniquely associated with the real wage. If commodities are in excess demand so that, given voluntary exchange, $y < y^D$, which in turn implies $x^{S'} < x^S$, at real wage w^* excess demand for labor will amount to quantity AF .

Second, with commodities in excess demand, the quantity of employment will generally be below the full employment level. The explanation of this apparent

¹⁸ For example, R. Vicker, a recent visitor to the Soviet Union, reports the effects of suppressed inflation upon output: "Goods produced for sale in state retail outlets are snapped up more and more quickly, and the remaining excess of income over things to spend it on dilutes the incentive of Soviet workers."

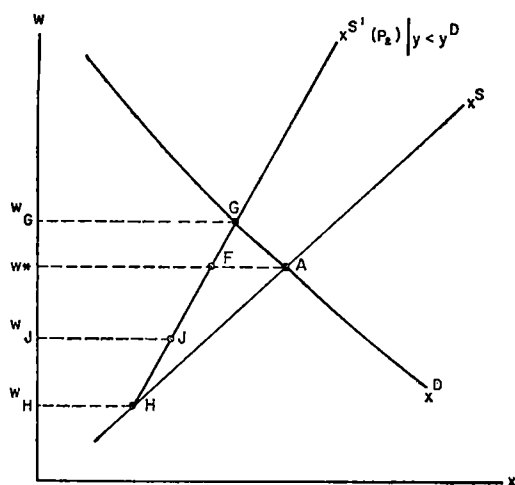


FIGURE 6. THE LABOR MARKET WITH EXCESS DEMAND FOR COMMODITIES

paradox, as indicated above, is twofold: 1) the quantity of employment can be no greater than the quantity supplied; and 2) when their consumption plans are frustrated households will generally substitute leisure and thus supply less labor at any given real wage. Notice that even if the real wage should rise sufficiently, i.e., to w_G , to eliminate the excess demand for labor, the level of employment would still be below that obtaining at general equilibrium.

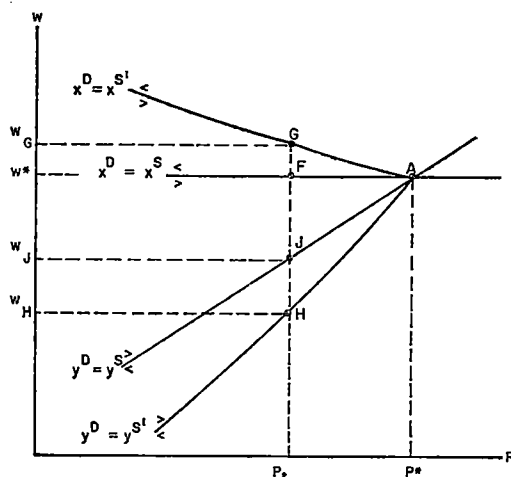


FIGURE 7. INTERACTION OF EXCESS DEMAND IN BOTH MARKETS

Finally, Figures 7 and 8, which are analogous to Figures 3 and 4, depict the interaction between the two markets with excess demand in both. Points A, F, G, H, and J in Figure 7 coincide with the same points in Figures 5 and 6. Figure 8 is drawn under the assumption that the existing wage-price vector is (w^*, P_2) , that is, that the economy is at point F. The details of the construction of these diagrams are left as an exercise for the reader.

VII. Summary

This paper describes the application of a general disequilibrium approach to familiar problems of macro-analysis. Some familiar results, such as the notion that insufficient commodity demand produces unemployment, are arrived at in a much more satisfactory manner than is possible under more conventional analysis. In addition, the specific inclusion of disequilibrium elements leads to some non-familiar results.

The impact of excess supply of commodities on labor demand removes the one-to-one classical relationship between real wage and employment. In a general disequilibrium situation, unemployment can coexist with "non-excessive" real wages, and a procyclical pattern of real wages is consistent with the theoretical model.

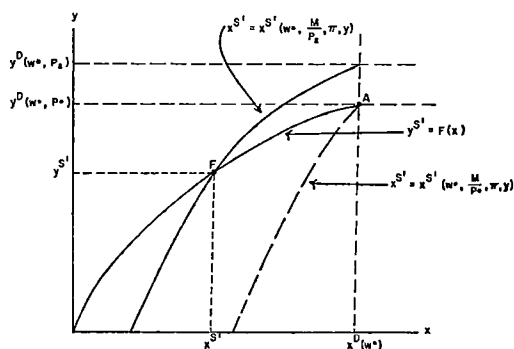


FIGURE 8. OUTPUT AND EMPLOYMENT WITH EXCESS DEMAND IN BOTH MARKETS

The disequilibrium analysis of the commodity market is formally parallel to the analysis of the labor market. The Keynesian consumption function emerges as a manifestation of the impact of excess labor supply on commodity demand. In this respect conventional macro-analysis is seen to be asymmetric. On the one hand, the disequilibrium impact of excess labor supply is implicitly recognized by entering income as a separate argument in the consumption function. However, on the other hand, the impact of excess commodity supply is neglected by adhering to the classical labor demand function which involves only the real wage. Because of this peculiar asymmetry, previous analyses of unemployment have had to rely on such contrived devices as a countercyclical pattern of real wages or fixed proportion production functions.

The framework for analyzing the excess supply, depression case is directly applicable to an analysis of sustained excess demand. The classical concept of forced saving is one aspect of the impact of excess commodity demand on household decision making. The forced saving solution is, however, incomplete, since labor supply would also react inversely to a prolonged frustration of commodity demand. To the extent that labor supply declines in response to excess commodity demand, increases in commodity demand lead to reduced employment, rather than to increased (forced) saving.

REFERENCES

- A. A. Alchian, "Information Costs, Pricing, and Resource Unemployment," *Western Econ. J.*, June 1969, 7, 109-28.
- K. J. Arrow, "Toward a Theory of Price Adjustment," in M. Abramowitz, ed., *The Allocation of Economic Resources*, Stanford 1959.
- R. J. Barro, "A Theory of Monopolistic Price Adjustment," read at Econometric Society Meetings, Detroit, Dec. 1970.
- , "A Theory of Optimal Adjustment," forthcoming 1971.
- R. G. Bodkin, "Real Wages and Cyclical Variations in Employment," *Can. J. Econ.*, Aug. 1969, 2, 353-74.
- R. Clower, "The Keynesian Counter-Revolution: A Theoretical Appraisal," in F. H. Hahn and F. P. R. Brechling, eds., *The Theory of Interest Rates*, London 1965.
- E. O. Edwards, "Classical and Keynesian Employment Theories: A Reconciliation," *Quart. J. Econ.*, Aug. 1959, 73, 407-28.
- H. I. Grossman, "Money, Interest, and Prices in Market Disequilibrium," *J. Polit. Econ.*, forthcoming 1971.
- J. Hicks, *Capital and Growth*, New York, 1965.
- J. M. Keynes, *The General Theory of Employment, Interest, and Money*, New York 1936.
- , "Relative Movements of Real Wages and Output," *Econ. J.*, Mar. 1939, 49, 34-51.
- E. Kuh, "Unemployment, Production Functions, and Effective Demand," *J. Polit. Econ.*, June 1966, 74, 238-49.
- A. Leijonhufvud, *On Keynesian Economics and the Economics of Keynes*, New York 1968.
- D. Patinkin, "Involuntary Unemployment and the Keynesian Supply Function," *Econ. J.*, Sept. 1949, 59, 360-83.
- , *Money, Interest, and Prices*, 1956: 2d ed. New York 1965.
- D. H. Robertson, *Banking Policy and the Price Level*, London 1926.
- R. M. Solow and J. E. Stiglitz, "Output, Employment, and Wages in the Short Run," *Quart. J. Econ.*, Nov. 1968, 82, 537-60.
- R. Vicker, "USSR: Rising Income, Black Markets," *The Wall Street Journal*, Apr. 21, 1970.

A Test for Relative Efficiency and Application to Indian Agriculture

By LAWRENCE J. LAU AND PAN A. YOTOPOULOS*

Economic efficiency is an elusive concept in which the economist, the engineer, and the policy maker all have great stakes.¹ The policy implications of economic efficiency permeate both the micro- and the macroeconomic level. Suppose, for example, that we can measure the efficiency of small and large farms. We can then determine by how much a given set of farms could be expected to increase its output through appropriate reorganization without absorbing additional resources in the aggregate.² We can also draw policy recommendations in connection with land ceilings, land redistribution, and land groupings under cooperative farming and other forms of agrarian organization.³

The difficulties with the existing approaches to efficiency are both conceptual and empirical. We will illustrate by pointing out the ambiguities of the conventional variants of efficiency, which we will classify

as economic efficiency, price or allocative efficiency and technical efficiency.⁴

The simplest—and most naïve—measure of economic efficiency is a partial productivity index, usually that of labor, although occasionally farm land (see Morton Paglin, 1965). This approach ignores the presence of other factors which affect average (and marginal) productivity.⁵ A more sophisticated approach constructs indexes of efficiency that consist of a weighted average of inputs (the weights being either relative prices or relative factor shares) which is compared to output. Such an index is basically an output-cost ratio (see Paglin 1965, Robert Bennett). This approach runs into the usual index number problems that have been so aptly summarized by Evsey Domar.

Price or allocative efficiency traditionally rests on an index of marginal product and opportunity cost.⁶ A number of problems arise in connection with this approach to price efficiency. First, it is an absolute concept which is of doubtful usefulness when one compares different groups of firms, even after allowing for differences in production functions and input prices. If among all inputs, the ratios of marginal products to opportunity costs are equal to one, a firm is price-efficient. Direct comparison among firms that satisfy

* The authors are, respectively, assistant professor, department of economics, and associate professor, Food Research Institute, Stanford University. An earlier draft of this paper benefited from the comments of A. S. Goldberger, B. F. Johnston, P. Zarembka, and the members of the Berkeley-Stanford Mathematical Economics and Econometrics Seminar. Mr. K. Somel provided able computational assistance. We also acknowledge financial support for Lau's participation by the National Science Foundation through Grant GS-2874 to Stanford University.

¹ For a discussion of the conceptual and theoretical complications that arise in connection with efficiency, see Margaret Lady Hall and Christopher Winsten.

² This is the approach of Harvey Leibenstein and William Comanor.

³ Most frequently studies of efficiency lead to this kind of microeconomic policy implication. Specific examples are supplied below.

⁴ In our own formulation below each of these terms will be given specific and rigorous conceptual and empirical content.

⁵ These other factors can be ignored with impunity only under certain special circumstances.

⁶ For examples, see Theodore Schultz, W. David Hopper, Yotopoulos (1968 a and b).

this equality to different degrees is almost impossible, as the literature on the second best has demonstrated. Second, it is a rather rigid concept that does not allow for possible differences in the initial endowment of fixed factors.

The conventional measurement of technical efficiency concentrates on the neutral displacement of the production function either between groups of firms or over time (see Irving Hoch, Yair Mundlak). Contrary to price efficiency which is purely a behavioral concept, technical efficiency is purely an engineering concept. It entirely abstracts from the effect of prices. An alternative approach to technical efficiency has been suggested by Michael Farrell. Under the assumption of constant returns to scale Farrell derives the "pessimistic" unit-isoquant, i.e., the isoquant which envelops the observations in the inputs-unit output space in such a way that no observation lies between the pessimistic isoquant and the origin. The index of efficiency is then constructed by measuring the deviation of a specific observation from the pessimistic isoquant. Besides also ignoring the effects of relative prices, the Farrell approach has the additional disadvantage that arises when one attempts to describe a stochastic universe by a deterministic process. The pessimistic isoquant is extremely sensitive to "outliers."⁷

The deficiencies of the existing approaches to measuring efficiency should dictate the minimum requirements that a new concept of relative *economic efficiency* should meet if it is to be at all useful. (i) It should account for firms that produce different quantities of output from a given set of measured inputs of production. This is the component of differences in *technical efficiency*. (ii) It should take into account

that different firms succeed to varying degrees in maximizing profits, i.e., in equating the value of the marginal product of each variable factor of production to its price. This is the component of *price efficiency*. (iii) The test should take into account that firms operate at different sets of market prices. The decision rule on profit maximization yields actual profits (as well as quantity of output supplied and quantities of variable inputs demanded) as a function, *inter alia*, of input prices. It is clear that two firms of equal technical efficiency which have successfully maximized profits would still have different value of profits as long as they face different prices.

The interrelationships of the concepts of technical efficiency, price efficiency and economic efficiency can be explained in an intuitive way. Consider two firms with production functions identical up to a neutral displacement parameter,

$$V^1 = A^1 F(X^1); \quad V^2 = A^2 F(X^2),$$

where V is the output, A the technical efficiency parameter, F the production function, X the vector of inputs employed, and the superscript denotes firm.

A firm is considered more technical-efficient than another, if, given the same quantities of measurable inputs, it consistently produces a larger output. Firm 1 is more technical-efficient than firm 2 if $A^1 > A^2$.

A firm is price-efficient if it maximizes profits, i.e., it equates the value of the marginal product of each variable input to its price. A firm which fails to maximize profits is, by definition, price inefficient. Consider now two complications in connection with the definition of price efficiency. First, assume that the prices of inputs are different for each firm. Firms now equate the value of the marginal product of each factor to its firm-specific opportunity cost. Second, firms may not maxi-

⁷ The approach has been modified to partly account for this shortcoming by Dennis Aigner and S. F. Chu and C. Peter Timmer.

mize profits. For such firms the usual marginal conditions do not hold. It is assumed that these firms equate the value of the marginal product of each factor to a constant (which may be firm- and factor-specific) proportion of the respective firm-specific factor prices, i.e., and for firm 1,

$$p \frac{\partial V^1}{\partial X_i^1} = k_i^1 c_i^1, \quad k_i^1 \geq 0$$

In this case k_i^1 indexes the decision rule that describes the firm's "profit-maximizing" behavior with respect to factor i . It encompasses perfect profit maximization as a special case when $k_i^1 = 1$ for all i . Now consider two price-inefficient firms of equal technical efficiency and facing identical output and input prices. The firm with the higher profits within a certain range of prices is considered the *relatively more price-efficient* firm (within the same range of prices).

Economic efficiency combines both technical and price efficiency. For this purpose consider two firms of *varying degrees of technical and price efficiency* but facing identical prices. The firm with the higher profits within a certain range of prices is considered the *relatively more economic-efficient* firm.

The concept of the profit function, as first introduced by D. L. McFadden, becomes operationally the ideal tool for our approach.⁸ In Section I, we develop the theory of the profit function in its general form, without introducing firm-specific (and input-specific) price efficiency decision rules or firm-specific technical efficiency parameters. These are introduced in Section II which formulates the test of relative economic efficiency for the general case. In Section III we make the

test of Section II operational by casting it in the framework of a Cobb-Douglas function.

Section IV is based on data from the *Farm Management Studies (The Studies)* of the Indian Ministry of Food and Agriculture. A number of researchers have used the same body of data to draw efficiency implications between small and large farms in India: A. K. Sen (1964, 1966); Paglin (1965); Bennett; G. S. Sahota; A. M. Khusro; to mention only a few. The findings are generally contradictory and inconclusive.⁹ This comes as small surprise, given the divergent and often ambiguous concepts of efficiency that the authors have used.

A warning is in order at this point. We share the reservations of the previous authors about the limitations and the reliability of the data of *The Studies*. The reader, therefore, is urged to interpret cautiously our finding that small farms are economically more efficient than large farms. We intend our empirical application as an illustration of a method of measuring relative efficiency. It is a method that is based on the precepts of economic theory, it is more general than the existing alternatives, it is operational and it is parsimonious from the point of view of data requirements. Needless to say, the usefulness of our test is not restricted to agriculture nor is it specific for comparing small and large farms. Actually much more important insights into the form of economic organization might be forthcoming if one compares different groupings, such as owners versus share tenants, leaseholders versus tenants, adopters of new varieties versus nonadopters. Similar ramifications can be suggested for other fields of economics.

⁸ Marc Nerlove first proposed a measurement of relative economic efficiency based on a profit function. However his approach is different from ours.

⁹ For a summary of this discussion see Jagdish Bhagwati and Sukhamoy Chakravarty, and Yotopoulos, Lau, and Kutlu Somel.

I. The Profit Function

Consider a firm with a production function with the usual neoclassical properties

$$(1) \quad V = F(X_1, \dots, X_m; Z_1, \dots, Z_n)$$

where V is output, X_i represents variable inputs, and Z_i represents fixed inputs of production. Profit (defined as current revenues less current total variable costs) can be written

$$P' = pF(X_1, \dots, X_m; Z_1, \dots, Z_n)$$

$$(2) \quad - \sum_{i=1}^m c'_i X_i$$

where P' is profit, p is the unit price of output, and c'_i is the unit price of the i th variable input. The fixed costs are ignored since, as it is well known, they do not affect the optimal combination of the variable inputs.

Assume that a firm maximizes profits given the levels of its technical efficiency and fixed inputs. The marginal productivity conditions for such a firm are

$$(3) \quad p \frac{\partial F(X; Z)}{\partial X_i} = c'_i, \quad i = 1, \dots, m$$

By using the price of the output as numeraire we may define $c_i \equiv c'_i/p$ as the normalized price of the i th input. We can then write (3) as

$$(4) \quad \frac{\partial F}{\partial X_i} = c_i, \quad i = 1, \dots, m$$

By similar deflation by the price of output we can rewrite (2) as (5) where we define P as the "Unit-Output-Price" profit (or UOP profit)

$$(5) \quad P = \frac{P'}{p} = F(X_1, \dots, X_m; Z_1, \dots, Z_n) - \sum_{i=1}^m c_i X_i$$

Equation (4) may be solved for the optimal quantities of variable inputs, denoted X_i^* 's, as functions of the normalized prices of the variable inputs and of the quantities of the fixed inputs,¹⁰

$$(6) \quad X_i^* = f_i(c, Z), \quad i = 1, \dots, m$$

where c and Z are the vectors of normalized input prices and quantities of fixed inputs, respectively.

By substitution of (6) into (2) we get the *profit function*,¹¹

$$(7) \quad \begin{aligned} \Pi &= p \left[F(X_1^*, \dots, X_m^*; Z_1, \dots, Z_n) \right. \\ &\quad \left. - \sum_{i=1}^m c'_i X_i^* \right] \\ &= G(p, c'_1, \dots, c'_m; Z_1, \dots, Z_n) \end{aligned}$$

The profit function gives the *maximized* value of the profit for each set of values $\{p; c'_1, \dots, c'_m; Z_1, \dots, Z_n\}$. Observe that the term within square brackets on the right-hand side of (7) is a function only of c and Z . Hence we can write

$$(8) \quad \Pi = pG^*(c_1, \dots, c_m; Z_1, \dots, Z_n)$$

The UOP profit function is therefore given by

$$(9) \quad \Pi^* = \frac{\Pi}{p} = G^*(c_1, \dots, c_m; Z_1, \dots, Z_n)$$

Observe also that maximization of profit in (2) is equivalent to maximization of UOP profit in (5) in that they yield identical values for the optimal X_i^* 's. Hence Π^* in (9) indeed gives the maximized value of UOP profit in (5). We employ the UOP profit function Π^* because it is easier to work with than Π . It is evident that given Π^* one can always find Π , and vice versa.

¹⁰ The unsubscripted variables X, Z, c, X^i, Z^i, c^i , and k^i are used to denote vectors. Superscripts, as above, denote firms.

¹¹ One should be careful to distinguish between profit as defined in (2) and the profit function in (7).

On the basis of a priori theoretical considerations we know that the *UOP* profit function is decreasing and convex in the normalized prices of variable inputs and increasing in quantities of fixed inputs. It follows also that the *UOP* profit function is increasing in the price of the output.

A set of dual transformation relations connects the production function and the profit function.¹² The most important one, from the point of view of our application here, is what is sometimes referred to as the Shephard-Uzawa-McFadden Lemma, as shown in equations (10) and (11).

$$(10) \quad X_i^* = - \frac{\partial \Pi^*(c, Z)}{\partial c_i}, \quad i = 1, \dots, m,$$

$$(11) \quad V^* = \Pi^*(c, Z) - \sum_{i=1}^m \frac{\partial \Pi^*(c, Z)}{\partial c_i} c_i$$

where V^* is the supply function.

At this point we should emphasize the advantages of working with the *UOP* profit function instead of the traditional production function. First, the Shephard-Uzawa-McFadden Lemma allows us to derive the firm's supply function, V^* , and the firm's factor demand functions, X_i^* 's, directly from the *UOP* profit function of (9) instead of solving equation (4) which involves the production function.¹³ Second, it is clear that the supply function and the factor demand functions may be obtained by simply starting with an *arbitrary* *UOP* profit function which is decreasing and convex in the normalized prices of the variable inputs and increasing in the fixed inputs. In addition, by duality, as McFadden has shown, there exists a one-to-one correspondence between the set of concave production functions and the set of convex

profit functions.¹⁴ Every concave production function has a dual which is a convex profit function, and vice versa.¹⁵ Hence, without loss of generality, one can consider for profit-maximizing, price-taking firms, only profit functions in the analysis of their behavior *without* an explicit specification of the corresponding production function. This provides a great deal of flexibility in empirical analysis. Third, by starting from a profit function, we are assured by duality that the resulting system of supply and factor demand functions is obtainable from the maximization of a concave production function subject to given fixed inputs and under competitive markets. Fourth, the profit function, the supply function, and the derived demand functions so obtained are functions only of the normalized input prices and the quantities of fixed inputs, variables that are normally considered to be determined independently of the firm's behavior. Econometrically, this implies that these variables are exogenous variables, and by estimating these functions we avoid the problem of simultaneous equations bias to the extent that it is present.

II. Relative Economic Efficiency

The discussion of the profit function in Section I is general. It does not consider differences in technical efficiency and differences in price efficiency that might exist between firms. The purpose of this section is to introduce such differences and to combine them into the concept of rela-

¹² These relations are given and proven in McFadden and Lau.

¹³ One practical advantage of using *UOP* profit functions as opposed to deriving the factor demand equations directly from equation (4) is that in many cases equation (4) cannot be solved in closed form.

¹⁴ There are additional regularity conditions on the production and profit functions which are spelled out in detail in McFadden. Since we are interested in the empirical application of profit functions, we will not be concerned with the finer details. It suffices to say that almost all continuous production functions in current use which are concave will give rise to a well-behaved profit function.

¹⁵ We rule out constant returns to scale in the variable factors, which, as is well known, would lead to indeterminate output and input levels. See Lau.

tive economic efficiency. Our approach is straightforward. Given comparable endowments, identical technology, and normalized input prices, the *UOP* profits of two firms should be identical if they have both maximized profits. To the extent that the one firm is more price-efficient, or more technically efficient, than the other, the *UOP* profits will differ even for the same normalized input prices and endowments of fixed inputs.

Let us represent the situation as follows. For each of two firms the production function is given by

$$(12) \quad \begin{aligned} V^1 &= A^1 F(X^1, Z^1); \\ V^2 &= A^2 F(X^2, Z^2) \end{aligned}$$

where superscripts identify firms. The marginal conditions are given by

$$(13) \quad \begin{aligned} \frac{\partial A^1 F(X^1, Z^1)}{\partial X_j^1} &= k_j^1 c_j^1 \\ \frac{\partial A^2 F(X^2, Z^2)}{\partial X_j^2} &= k_j^2 c_j^2 \\ k_j^1 &\geq 0, k_j^2 \geq 0, j = 1, \dots, m \end{aligned}$$

At this point it is useful to reiterate the basic differences in approach that equations (12) and (13) introduce, as compared to Section I. The formulation of Section I was general while now it becomes firm-specific. We can talk about relative efficiency only by comparing two or more firms. We allow for neutral differences in the production functions in terms of the firm-specific technical efficiency parameters, A^1 and A^2 . They represent differences in environmental factors, in managerial ability and in other nonmeasurable fixed factors of production. If the two firms are equally technical-efficient, $A^1 = A^2$. Furthermore, we now allow for a firm to be unsuccessful in its attempts to equate values of the marginal products of its inputs to their respective normalized prices.

This is introduced through the firm-specific and variable input-specific k 's.¹⁶ If, and only if, two firms are equally price-efficient with respect to all variable inputs, then $k_i^1 = k_i^2$, $i = 1, \dots, m$. We have defined economic efficiency to encompass both technical and price efficiency. In terms of our notation, therefore, the null hypothesis of equal relative economic efficiency for firm 1 and firm 2 implies that $A^1 = A^2$ and $k^1 = k^2$. The purpose of this section, therefore, is to develop a method to enable us to make this comparison.

In our formulation, the k 's reflect a general systematic rule of behavior—a decision rule that gives the profit-maximizing marginal productivity conditions as a special case. That the decision rule for the firm consists of equating the marginal product to a constant times the normalized price of each input may be rationalized as follows: i) Consistent over- or under-valuation of the opportunity costs of the resources by the firm; ii) Satisficing behavior; iii) Divergence of expected and actual normalized prices; iv) Divergence of the subjective probability distribution of the normalized prices from the objective distribution of normalized prices; v) The elements of k^i may be interpreted as the first-order coefficients of a Taylor's series expansion of arbitrary decision rules of the type

$$\frac{\partial F}{\partial X_j^i} = f_j^i(c_j^i), \quad i = 1, 2; j = 1, \dots, m$$

where $f_j^i(0) = 0$ and $f_j^{i'}(c_j^i) \geq 0$. A wide class of decision rules may be encompassed under v). Observe that the right-hand sides of equation (13) may be interpreted as the "effective" prices facing the two firms. The behavior of the two firms can

¹⁶ Of course, if a firm is perfectly successful in equalizing the normalized price of an input i to its opportunity cost, k_i assumes the value of one for that specific input.

then be viewed as profit-maximization subject to these effective prices and can be represented by the *behavioral UOP* profit function.

Let $G^*(c, Z)$ be the *UOP* profit function corresponding to $F(X, Z)$. By a well-known theorem proved in McFadden, the *UOP* profit function corresponding to a production function

$$(14) \quad \begin{aligned} V &= AF(X, Z) \quad \text{is} \\ \Pi^* &= AG^*(c/A, Z) \end{aligned}$$

Recall that the $k_j^i c_j^i$'s may be interpreted as the effective prices. Thus we may write for the behavioral *UOP* profit functions of the two firms, respectively,

$$(15) \quad \begin{aligned} \Pi_b^1 &= A^1 G^*(k_1^1 c_1^1 / A^1, \dots, k_m^1 c_m^1 / A^1; \\ &\quad Z_1^1, \dots, Z_n^1) \\ \Pi_b^2 &= A^2 G^*(k_1^2 c_1^2 / A^2, \dots, k_m^2 c_m^2 / A^2; \\ &\quad Z_1^2, \dots, Z_n^2) \end{aligned}$$

As in the previous section, the demand functions are given by the Shephard-Uzawa-McFadden Lemma. We now, however, differentiate the behavioral *UOP* profit functions with respect to the effective prices $k_j^1 c_j^1$'s and $k_j^2 c_j^2$'s. We write¹⁷

$$(16) \quad \begin{aligned} X_j^i &= -A^i \frac{\partial G^*(k^i c^i / A^i; Z^i)}{\partial k_j^i c_j^i} \\ &= \frac{-A^i}{k_j^i} \frac{\partial G^*(k^i c^i / A^i; Z^i)}{\partial c_j^i}, \\ &\quad i = 1, 2; j = 1, \dots, m \end{aligned}$$

By correspondence from (11) the supply functions are now given by

$$\begin{aligned} V^i &= A^i G^*(k^i c^i / A^i; Z^i) \\ &- A^i \sum_{j=1}^m \frac{k_j^i}{k_j^i c_j^i} \frac{\partial G^*(k^i c^i / A^i; Z^i)}{\partial k_j^i c_j^i} \end{aligned}$$

¹⁷ To simplify notation we omitted the asterisks from the demand and supply functions.

$$(17) \quad \begin{aligned} &= A^i G^*(k^i c^i / A^i; Z^i) \\ &- A^i \sum_{j=1}^m c_j^i \frac{\partial G^*(k^i c^i / A^i; Z^i)}{\partial c_j^i}, \\ &\quad i = 1, 2 \end{aligned}$$

It should be emphasized at this point that X_j^i and V^i as given in (16) and (17) are the actual quantities of inputs demanded and output supplied by firm i given the firm-specific A^i and k^i . When appropriate functional forms are specified for G , statistical tests can be devised to test the null hypothesis of equal economic efficiency, i.e., $A^1 = A^2$ and $k^1 = k^2$, although not all of the parameters may be independently identified and estimated.

An alternative approach to looking at the demand and supply functions is to examine the *actual UOP* profit function. From (16) and (17) we can obtain the actual *UOP* profit functions by using equation (5),

$$(18) \quad \begin{aligned} \Pi_a^i &= V^i - \sum_{j=1}^m c_j^i X_j^i \\ &= A^i G^*(k^i c^i / A^i; Z^i) \\ &+ A^i \sum_{j=1}^m \frac{(1 - k_j^i) c_j^i}{k_j^i} \\ &\quad \cdot \frac{\partial G^*(k^i c^i / A^i; Z^i)}{\partial c_j^i}, \quad i = 1, 2 \end{aligned}$$

Observe that i) $\partial \Pi_a^i / \partial A^i > 0$, i.e., actual profit always increases with the level of technical efficiency for given normalized input prices and k^i ; ii) When $k_j^i = 1$ for $j = 1, \dots, m$, i.e., the firm is a true profit maximizer, the actual and behavioral *UOP* profit functions coincide; iii) When $A^1 = A^2$ and $k^1 = k^2$, the actual *UOP* functions of the two firms coincide with each other. Therefore one can also test the null hypothesis of equal relative economic efficiency by comparing the actual *UOP* profit functions of the two firms when appropriate functional forms are specified

for G . This is the approach that will be employed in our empirical analysis.¹⁸

An additional test becomes relevant if we reject the joint hypothesis that $(A^1, k^1) = (A^2, k^2)$. In this case an overall indication of the relative efficiency between the two firms within a specified range of normalized prices for variable inputs may be obtained by comparing the actual values of the UOP profit functions within this range. If $\Pi_a^1 \geq \Pi_a^2$ for all normalized prices within a specified range, then clearly, the first firm is relatively more efficient within the price range. If some knowledge on the probability distribution of the future prices is available, a choice may be made as to the relative efficiency of the two firms.

One can also test the hypothesis that the fixed inputs command equal rent on the two firms by computing the first derivatives of the actual UOP profit functions with respect to the fixed inputs and testing for their equality. This may have important implications for the optimal form of economic organization in terms of the distribution of fixed inputs.

III. The Formulation of the Cobb-Douglas Case

In this section we proceed to specify the appropriate functional form of the profit function and formulate empirically the test of relative economic efficiency. For this purpose one can start from a Cobb-

Douglas, or for that matter, from any other form of a function. We cast our analysis in terms of the Cobb-Douglas function because it appears superior through tests of alternative functional forms.¹⁹

A Cobb-Douglas production function with decreasing returns in the m variable inputs and with n fixed inputs is given by²⁰

$$V = A \left(\prod_{i=1}^m X_i^{\alpha_i} \right) \left(\prod_{j=1}^n Z_j^{\beta_j} \right)$$

where $\mu = \sum_{i=1}^m \alpha_i < 1$

The UOP profit function is given by

$$\begin{aligned} \Pi^* = & A^{(1-\mu)^{-1}} (1 - \mu) \\ & \cdot \left(\prod_{i=1}^m (c_i / \alpha_i)^{-\alpha_i (1-\mu)^{-1}} \right) \\ & \cdot \left(\prod_{j=1}^n Z_j^{\beta_j (1-\mu)^{-1}} \right) \end{aligned} \quad (19)$$

By direct computation, the actual UOP profit functions and the demand functions for this Cobb-Douglas production function are given in equations (20) and (21). It is clear that the actual UOP profit functions of the two firms differ by a constant factor, which is a function of the k_j 's and A 's. In addition, all the demand functions differ by constant factors. A test of equal economic efficiency will be based on the

¹⁸ Note that by the profit identity, one of the system of profit, supply and demand functions is redundant and should be ignored in the actual estimation of the system. Otherwise the system variance-covariance matrix will be singular.

¹⁹ These tests are presented in Yotopoulos, Lau, and Somel.

²⁰ The value of $\mu < 1$ is required since constant or increasing returns in the variable inputs are inconsistent with profit maximization.

$$\begin{aligned} \Pi_a^i = & \left[(A^i)^{(1-\mu)^{-1}} (1 - \sum_{j=1}^m \alpha_j / k_j^i) \right] \left[\prod_{j=1}^m (k_j^i)^{-\alpha_j (1-\mu)^{-1}} \right] \left[\prod_{j=1}^m \alpha_j^{-\alpha_j (1-\mu)^{-1}} \right] \\ & \cdot \left[\prod_{j=1}^m (c_j^i)^{-\alpha_j (1-\mu)^{-1}} \right] \left[\prod_{j=1}^n (Z_j^i)^{\beta_j (1-\mu)^{-1}} \right], \quad i = 1, 2 \end{aligned} \quad (20)$$

$$(21) \quad X_l^i = (A^i)^{(1-\mu)^{-1}} (\alpha_l/k_l c_l^i) \left[\prod_{j=1}^m (k_j^i)^{-\alpha_j(1-\mu)^{-1}} \right] \left[\prod_{j=1}^m \alpha_j^{-\alpha_j(1-\mu)^{-1}} \right] \\ \cdot \left[\prod_{j=1}^m (c_j^i)^{-\alpha_j(1-\mu)^{-1}} \right] \left[\prod_{j=1}^n (Z_j^i)^{\beta_j(1-\mu)^{-1}} \right], \quad i = 1, 2; \quad l = 1, \dots, m$$

null hypothesis that all the constant factors of difference are ones.

Observe that the terms in the first three brackets of equation (20) involve constants. We thus define equation (22).

$$(22) \quad A_*^i \equiv (A^i)^{(1-\mu)^{-1}} \left(1 - \sum_{j=1}^m \alpha_j/k_j^i \right) \\ \left[\prod_{j=1}^m (k_j^i)^{-\alpha_j(1-\mu)^{-1}} \right] \left[\prod_{j=1}^m \alpha_j^{-\alpha_j(1-\mu)^{-1}} \right], \\ i = 1, 2$$

Then the actual *UOP* profit functions are given by

$$(23) \quad \Pi_a^i = (A_*^i) \left[\prod_{j=1}^m (c_j^i)^{-\alpha_j(1-\mu)^{-1}} \right] \\ \cdot \left[\prod_{j=1}^n (Z_j^i)^{-\beta_j(1-\mu)^{-1}} \right], \\ i = 1, 2$$

By writing A_*^2 and A_*^1 for firm 2 and firm 1, respectively, and taking the ratio of the constant terms we have

$$(24) \quad \frac{A_*^2}{A_*^1} = \left[\frac{A^2}{A^1} \right]^{(1-\mu)^{-1}} \frac{\left(1 - \sum_{j=1}^m \alpha_j/k_j^2 \right)}{\left(1 - \sum_{j=1}^m \alpha_j/k_j^1 \right)} \\ \cdot \left[\prod_{j=1}^m \left[\frac{k_j^2}{k_j^1} \right]^{-\alpha_j(1-\mu)^{-1}} \right]$$

Thus one may write, from equation (20)

$$(25) \quad \Pi_a^1 = A_*^1 \left[\prod_{j=1}^m (c_j^1)^{-\alpha_j(1-\mu)^{-1}} \right] \\ \cdot \left[\prod_{j=1}^n (Z_j^1)^{\beta_j(1-\mu)^{-1}} \right]$$

$$(26) \quad \Pi_a^2 = A_*^1 (A_*^2/A_*^1) \left[\prod_{j=1}^m (c_j^2)^{-\alpha_j(1-\mu)^{-1}} \right] \\ \cdot \left[\prod_{j=1}^n (Z_j^2)^{\beta_j(1-\mu)^{-1}} \right]$$

Further defining

$$(27) \quad \alpha_j^* \equiv -\alpha_j(1-\mu)^{-1};$$

and

$$(28) \quad \beta_j^* \equiv \beta_j(1-\mu)^{-1}$$

and taking natural logarithms of equations (25) and (26), we have

$$(29) \quad \ln \Pi_a^1 = \ln A_*^1 + \sum_{j=1}^m \alpha_j^* \ln c_j^1 \\ + \sum_{j=1}^n \beta_j^* \ln Z_j^1,$$

$$(30) \quad \ln \Pi_a^2 = \ln A_*^1 + \ln \frac{A_*^2}{A_*^1} + \sum_{j=1}^m \alpha_j^* \ln c_j^2 \\ + \sum_{j=1}^n \beta_j^* \ln Z_j^2$$

We note that if $A^1 = A^2$ and $k^1 = k^2$, then $A_*^1 = A_*^2$ and the two functions Π_a^1 and Π_a^2 ($\ln \Pi_a^1$ and $\ln \Pi_a^2$) should be identical. This implies that $\ln A_*^2/A_*^1 = 0$. We can

therefore test the equal relative efficiency hypothesis by utilizing a firm dummy variable in the logarithmic *UOP* profit function and examining if its value is equal to zero. It should be noted that for the Cobb-Douglas production function case, differences in technical efficiency and relative differences in price efficiency cannot be separately identified from the actual *UOP* profit functions.

IV. Empirical Implementation and Statistical Results

In this section we use data from *The Studies* of the Indian Ministry of Food and Agriculture to estimate the *UOP* profit functions for the small and large farms and to apply the test of equal economic efficiency for the two groups. *The Studies*, which have proven to be a bountiful data source for many researchers,²¹ are based on cost-accounting records of 2,962 holdings in the six states of India and cover the three-year period, 1955-57.²² All the data are, however, reported only in terms of averages of farms of a given size for each state. From the available raw data we proceed to specify as follows the variables of our analysis.

Output is given in terms of revenue V per farm in rupees; land T represents cultivable land per farm in acres,²³ and capital K is defined in terms of interest charges paid or imputed on the quantity of fixed capital per farm.²⁴ Labor is given

²¹ Besides the literature based on *The Studies* that is surveyed in Bhagwati and Chakravarty (especially pp. 40 ff.), one should also notice Paglin, Sahota and Yotopoulos, Lau and Somel.

²² For this analysis we utilize data from the following states and years: West Bengal, Madras, Uttar Pradesh, Punjab, 1955-56; Madhya Pradesh, 1956-57. The latter is chosen because the 1955-56 report of *The Studies* for Madhya Pradesh does not contain comparable information as the others.

²³ It is assumed that the land input is homogeneous at least within states across farm sizes. This hypothesis was tested in Yotopoulos, Lau, and Somel.

²⁴ This definition of the capital concept is especially

in terms of labor days employed per farm as well as in terms of a labor cost per farm concept (i.e., cost of hired labor plus imputed cost of family labor). By dividing the latter labor concept through by the former we define the money wage rate per day, w' . Only three inputs are distinguished: labor, capital, and land. We treat labor as the variable input of production and land and capital as fixed inputs.²⁵ It appears reasonable, from both institutional reasons and from the periodic nature of the agricultural technology, that the latter may be considered as fixed inputs in the short run. Finally, from the revenue we subtract the total cost of variable inputs per farm, i.e., the wage bill, in order to define the profit variable, Π . It should be recalled that in the *UOP* profit function formulation of the preceding sections both Π^* and w are expressed in real terms. Unavailability of the prices for deflation poses a problem that will be discussed below.

For the Cobb-Douglas case, the profit function is given by (29) as

$$(31) \quad \ln \Pi_a^1 = \ln A_*^1 + \alpha_1^* \ln w \\ + \beta_1^* \ln K + \beta_2^* \ln T$$

disturbing. Inasmuch as the interest rate used in the imputation—3 percent—is uniform throughout the states and the years, the true quantity of fixed capital will be proportional to our measure. This implicitly assumes that the flow of capital services as a ratio of the stock of capital is constant across farms. Such assumptions, as demonstrated by Yotopoulos (1967, 1968a), may lead to unreliable estimates of the coefficient of capital in a production function formulation. It appears that this may be the case with our estimated capital coefficient.

²⁵ Total other costs (i.e., costs other than labor costs, interest on fixed capital and land rent) should also be treated as a variable input of production. This is impossible in our profit function formulation due to the fact that we lack the "price" of other costs which is necessary for the *UOP* profit function. To the extent that the price of other costs varies only across states, its effect is captured by the state dummies. An alternative rationalization is that the other costs are employed in fixed proportions to output.

$$(32) \quad \ln \Pi_a^2 = \ln A_*^1 + \ln (A_*^2/A_*^1) \\ + \alpha_1^* \ln w + \beta_1^* \ln K + \beta_2^* \ln T$$

where Π_a^i is actual *UOP* profit (total revenue less total variable cost, divided by the price of output), w is normalized wage rate, K is interest on fixed capital, and T is cultivable land. A maintained hypothesis is that the production function is identical on large and small farms up to a neutral efficiency parameter. This implies that the coefficients corresponding to $\ln w$, $\ln K$, and $\ln T$ are identical for large and small farms. A problem arises at this point. Our formulation of the *UOP* profit function is in terms of normalized input prices. However, in our empirical application these normalized input prices are not available since the data on money prices of output are rather poor. Fortunately, we note that equations (31) and (32) may be rewritten

$$\ln \Pi_a^1 = \ln \Pi'^1 - \ln p \\ = \ln A_*^1 + \alpha_1^* \ln w' - \alpha_1^* \ln p \\ + \beta_1^* \ln K + \beta_2^* \ln T$$

or

$$\ln \Pi^1 = \ln A_*^1 + (1 - \alpha_1^*) \ln p \\ + \alpha_1^* \ln w' + \beta_1^* \ln K \\ + \beta_2^* \ln T \\ \ln \Pi^2 = \ln A_*^1 + (1 - \alpha_1^*) \ln p \\ + \ln \left(\frac{A_*^2}{A_*^1} \right) + \alpha_1^* \ln w' \\ + \beta_1^* \ln K + \beta_2^* \ln T$$

where Π'^i is money profit in rupees, w' , is the money wage rate in rupees per day, and p is the price of the output in rupees.

If the prices of outputs differ only across states, then one can insert state dummy variables to capture the effect of differences due to $(\ln A_* + (1 - \alpha_1^*) \ln p)$. This also allows for interstate differences in the efficiency parameter in A_* . Hence our final estimating equation consists of

$$\ln \Pi = \alpha_0 + S + \sum_{i=1}^4 \delta_i^* D_i + \alpha_1^* \ln w' \\ (33) \quad + \beta_1^* \ln K + \beta_2^* \ln T$$

where Π is farm profit in rupees (excluding interest on capital and land rent), w' is money wage rate, K is interest on fixed capital, T is cultivable land, the D_i 's are state dummy variables and S is a dummy variable with value 1 for large farms and 0 for small farms. Large farms are defined as those with cultivable land greater than ten acres per farm.

A remark about the stochastic specification of the model is appropriate at this point. Not much is known about how disturbance terms in general should be introduced into economic relationships although Hoch, Mundlak and Hoch, and subsequently Zellner, Kmenta and Drèze have proposed one possible assumption that is workable in the Cobb-Douglas case. Here we assume that the error in the profits is due to climatic variations, divergence of the expected output price from the realized output price, imperfect knowledge of the technical efficiency parameter of the farm, and differences in technical efficiency among farms within the same size class. The demand functions are exact, or, in any case, if they are subject to error, the errors are uncorrelated with the errors of the logarithmic profit function, which are assumed to have zero expectation. Hence one can estimate the natural logarithms of the profit function alone with the least squares estimator, which in this case turns out to be minimum

variance, linear and unbiased. This specification is similar to the one used by Marc Nerlove in his pioneering study of empirical cost functions.

The results of the estimation are presented in Table 1. The F -value indicates that the hypothesis that all coefficients other than α_0 are zeroes should be rejected. The coefficient of the wage rate is negative while the coefficient of land is positive, in accord with a priori economic theory: the UOP profit function is decreasing in w' and increasing in T . The negative coefficient of capital can only be attributed to the misspecification of this variable that is due to the implicit assumption of proportionality between capital service flow and capital stock (Yotopoulos, 1967, 1968a). In addition, the second derivative of the UOP profit function with respect to the wage rate is

$$\begin{aligned}\frac{\partial^2 \Pi^*}{\partial w^2} &= \frac{\alpha_1^* \Pi^*}{w^2} - \frac{\alpha_1^* \Pi^*}{w^2} \\ &= \alpha_1^* (\alpha_1^* - 1) \frac{\Pi^*}{w^2} \\ &= -2.141 (-3.141) \frac{\Pi^*}{w^2} \geq 0\end{aligned}$$

as $\Pi > 0$ and hence $\Pi^* > 0$ for our whole sample. This also confirms the convexity assumption of the UOP profit function.

The estimates of Table 1 imply, by (27) and (28), estimates of the input elasticities of the production function. These are presented in Table 2. The elasticities appear reasonable by comparison with other available estimates of Cobb-Douglas agricultural production functions for India and other parts of the world.

We note that the capital coefficient as estimated directly also has a negative sign. Finally, the sum of elasticities obtained from the indirect estimates is somewhat larger than that obtained from the direct

TABLE 1—COBB-DOUGLAS PROFIT FUNCTION AND RELATED STATISTICS

Parameter	All Farms ($n=34$)
α_0	4.582 (0.548)
S	-0.567 (0.253)
δ_1^*	1.614* (0.549)
δ_2^*	-1.359* (1.274)
δ_3^*	-0.588* (0.485)
δ_4^*	0.296 (0.715)
α_i^*	-2.141** (1.200)
β_1^*	-0.588 (0.274)
β_2^*	1.797 (0.233)
$\hat{\sigma}^2$	0.185
\bar{R}^2	0.896
F -Statistic	36.4

Source: *Farm Management Studies*

Notes: The estimating equation is

$$\ln \Pi = \alpha_0 + S + \sum_{i=1}^4 \delta_i^* D_i + \alpha_1^* \ln w' + \beta_1^* \ln K + \beta_2^* \ln T$$

where

Π = profit, i.e., total revenue less total variable costs

w' = the money wage rate

S = dummy variable for farm size with value of one for large farms (greater than ten acres) and zero for small farms (less than ten acres)

D_i = regional dummy variable with D_1, D_2, D_3, D_4 taking the value of one for West Bengal, Madras, Madhya Pradesh, and Uttar Pradesh and zero elsewhere, respectively

K = interest on fixed capital

T = cultivable land in acres

$\hat{\sigma}^2$ = the estimate of the variance of the error in the equation

* Starred coefficients are not significantly different from zero at a probability level ≥ 95 percent

** Double-starred coefficients are not significantly different from zero at a probability level ≥ 95 percent; but they are significantly different from zero at a probability level ≥ 90 percent.

All other coefficients are significantly different from zero at a probability level ≥ 95 percent.

Two-tail test applies to the coefficients of the dummy variables; one-tail test to all other variables.

The standard errors of the estimated parameters are given in parentheses.

TABLE 2—COMPARISON OF DIRECT AND INDIRECT ESTIMATES OF INPUT ELASTICITIES OF THE PRODUCTION FUNCTION

Parameters	Direct Estimates	Indirect Estimates
α_1	.606	.682
β_1	-.103	-.187
β_2	.365	.572
$(\alpha_1 + \beta_1 + \beta_2)$	0.868	1.067

Source: *Farm Management Studies*.

Notes: The direct estimates are obtained by ordinary least squares regression of the natural logarithm of output on the natural logarithms of the three inputs, the farm size dummy, the four state dummies and the constant term. The indirect estimates of the parameters are derived from equations (27) and (28). Other notations as in Table 1.

estimates. It is slightly larger than one.²⁶ One may also add that the estimates obtained by fitting the profit function are statistically consistent, as opposed to those obtained directly from the production function by ordinary least squares, which are in general inconsistent because of the existence of simultaneous equations bias.

As we have indicated in the previous section the hypothesis of relative efficiency can be cast in terms of the constant term by which the two profit functions, one for small and one for large farms, differ. The null hypothesis is that the constant factor is equal to one. Furthermore, if one takes natural logarithms before estimating the profit function, the constant term becomes the coefficient of a dummy variable that differentiates the two groups of farms and the test becomes that the coefficient of the dummy variable is not significantly different from zero. Our results, therefore, reject the hypothesis of equal efficiency between the two groups. Furthermore, the sign of the dummy variable indicates that small farms are more profitable, i.e., more efficient, at all observed prices of the variable

input, given the distribution of the fixed factors of production.

Given the actual profit functions for the two groups of farms, one can estimate the rate of return on the fixed inputs by computing the partial derivatives of the *UOP* profit function with respect to both capital and land. Suppressing interstate differences, one has

$$(34) \quad \frac{\partial \Pi}{\partial K} = \beta_1^* \frac{\Pi}{K} = -0.588 \frac{\Pi}{K}$$

$$(35) \quad \frac{\partial \Pi}{\partial T} = \beta_2^* \frac{\Pi}{T} = 1.797 \frac{\Pi}{T}$$

These rates of returns are computed at the geometric mean of the large and small farms separately, and reported in Table 3. It is seen that both the rates of return on fixed capital and on land are larger on the small farms at the existing set of normalized prices faced by these farms.

V. Summary and Conclusion

In our formulation of the test of equal relative economic efficiency we use McFadden's profit function, which expresses a firm's maximized profit as a function of the prices of output and variable inputs of production and of the quantities of the

TABLE 3—COMPARISON OF THE RATES OF RETURN ON FIXED CAPITAL AND LAND BETWEEN LARGE AND SMALL FARMS

	Large Farms	Small Farms
Geometric Means		
$\bar{\Pi}$ (rupees)	2,184.62	493.90
\bar{K} (rupees)	51.35	22.89
\bar{T} (acres)	23.81	3.99
Rates of Return		
$\frac{\partial \Pi}{\partial K}$ (rupees per rupee)	-25.02	-12.69
$\frac{\partial \Pi}{\partial T}$ (rupees per acre)	164.88	222.44

Source: *Farm Management Studies*.

²⁶ An exact linear restriction is available for the testing of the hypothesis of constant returns to scale within the profit function framework. See Lau, also Lau and Yotopoulos.

fixed factors. In the Cobb-Douglas formulation the comparison of relative efficiency of two groups of firms is simply made by examining the coefficient of the group dummy variable.

A crucial feature of the profit function analysis is that it assumes firms behave according to certain decision rules, which include the profit maximization rules, given the price regime for output and variable inputs and given the quantities of their fixed factors of production. For the purposes of this paper, the existence of these systematic decision rules is a maintained hypothesis. It can, however, be tested directly within the framework developed by Wise and Yotopoulos.

The conclusion of the test of relative economic efficiency is in favor of the small farms (i.e., farms of less than ten acres). It appears that, given the fixed factors of production (land and fixed capital) and within the ranges of the observed prices of output and variable inputs (labor), the small farms of the sample of *The Studies* have higher actual profits. In the context of our analysis, this finding means that the small farms attain higher levels of price efficiency (i.e., of optimal price behavior) and/or they operate at higher levels of technical efficiency. This finding, should it be confirmed by similar tests with other sets of data, may imply that in agriculture the supervisory role of the owner-manager of the farm may be crucial for attaining high levels of economic efficiency. Within the context of this hypothesis, the test would draw the limits of the supervisory capacity of the manager at the ten acre farms.

REFERENCES

- D. J. Aigner and S. F. Chu, "On Estimating the Industry Production Function," *Amer. Econ. Rev.*, Sept. 1968, 58, 826-39.
- R. L. Bennett, "Surplus Agricultural Labor and Development—Facts and Theories: Comment," *Amer. Econ. Rev.*, Mar. 1967, 57, 194-202.
- J. N. Bhagwati and S. Chakravarty, "Contributions to Indian Economic Analysis: A Survey," *Amer. Econ. Rev. Supp.*, Sept. 1969, 59, 1-73.
- W. S. Comanor and H. Leibenstein, "Allocative Efficiency, X-Efficiency and the Measurement of Welfare Losses," *Economica*, Aug. 1969, 36, 304-09.
- E. Domar, "On Total Productivity and All That," *J. Polit. Econ.*, Dec. 1962, 70, 597-608.
- M. J. Farrell, "The Measurement of Productive Efficiency," *J. Roy. Statist. Soc.*, 1957, Part 3, 120, 253-82.
- M. L. Hall and C. B. Winsten, "The Ambiguous Notion of Efficiency," *Econ. J.*, Mar. 1959, 69, 71-86.
- I. Hoch, "Estimation of Production Function Parameters and Testing for Efficiency," *Econometrica*, 1955, 23, 325-26.
- , "Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function," *Econometrica*, Oct. 1958, 26, 556-78.
- W. D. Hopper, "Allocation Efficiency in a Traditional Indian Agriculture," *J. Farm Econ.*, Aug. 1965, 47, 611-24.
- A. M. Khusro, "Returns to Scale in Indian Agriculture," *Indian J. Agr. Econ.*, Oct. 1964, 19, 51-80.
- L. J. Lau, "Applications of Profit Functions," in D. L. McFadden, ed., *The Econometric Approach to Production Theory*, Amsterdam 1971, forthcoming.
- and P. A. Yotopoulos, "Profit, Supply and Factor Demand Functions with Application to Indian Agriculture," Memorandum No. 98, Research Center in Economic Growth, Stanford Univ., 1970.
- H. Leibenstein, "Allocative Efficiency vs. 'X-Efficiency'," *Amer. Econ. Rev.*, June 1966, 56, 392-415.
- D. L. McFadden, "Cost, Revenue, and Profit Functions," in D. L. McFadden, ed., *The Econometric Approach to Production Theory*, Amsterdam 1971, forthcoming.
- Y. Mundlak, "Empirical Production Function Free of Management Bias," *J. Farm Econ.*, Feb. 1961, 43, 44-56.

- and I. Hoch, "Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions," *Econometrica*, Oct. 1965, 33, 814-28.
- M. Nerlove, "Returns to Scale in Electricity Supply," in C. F. Christ et al., eds., *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, Stanford 1960, 167-98.
- , *Estimation and Identification of Cobb-Douglas Production Functions*, Chicago, 1965.
- M. Paglin, "'Surplus' Agricultural Labor and Development—Facts and Theories," *Amer. Econ. Rev.*, Sept. 1965, 55, 815-34.
- , "Surplus Agricultural Labor and Development—Facts and Theories: Reply," *Amer. Econ. Rev.*, Mar. 1967, 57, 202-09.
- G. S. Sahota, "Efficiency of Resource Allocation in Indian Agriculture," *Amer. J. Agr. Econ.*, Aug. 1968, 50, 584-605.
- T. W. Schultz, *Transforming Traditional Agriculture*, New Haven 1964.
- A. K. Sen, "Size of Holdings and Productivity," *Econ. Weekly*, Feb. 1964, 16, 323-26.
- , "Peasants and Dualism With or Without Surplus Labor," *J. Polit. Econ.*, Oct. 1966, 74, 425-50.
- R. W. Shephard, *Cost and Production Functions*, Princeton 1953.
- C. P. Timmer, "On Measuring Technical Efficiency," unpublished doctoral dissertation, Harvard Univ. 1969.
- H. Uzawa, "Duality Principles in the Theory of Cost and Production," *Int. Econ. Rev.*, May 1964, 5, 216-20.
- J. Wise and P. A. Yotopoulos, "The Empirical Content of Economic Rationality: A Test for a Less Developed Economy," *J. Polit. Econ.*, Nov. 1969, 77, 976-1004.
- P. A. Yotopoulos, (1968a) *Allocative Efficiency in Economic Development: A Cross Section Analysis of Epirus Farming*, Center of Planning and Economic Research, Athens 1968.
- , (1968b) "On the Efficiency of Resource Utilization in Subsistence Agriculture," *Food Res. Inst. Stud. Agr. Econ., Trade, Devel.*, 1968, 8, 125-35.
- , "From Stock to Flow Capital Inputs for Agricultural Production Functions: A Microanalytic Approach," *J. Farm Econ.*, May 1967, 49, 476-91.
- and L. J. Lau, "A Simultaneous Equation Approach to Relative Economic Efficiency," Memorandum No. 104, Research Center in Economic Growth, Stanford University 1970.
- , ———, and K. Somel, "Labor Intensity and Relative Efficiency in Indian Agriculture," *Food Res. Inst. Stud. Agr. Econ., Trade, Devel.*, 1970, 9, 43-55.
- A. Zellner, J. Kmenta and J. Drèze, "Specification and Estimation of Cobb-Douglas Production Function Models," *Econometrica*, Oct. 1966, 34, 784-95.
- The Government of India, Ministry of Food and Agriculture, *Studies in the Economics of Farm Management*, Delhi 1957-62. Reports for the year 1955-1956: Madras, Punjab, Uttar Pradesh, West Bengal; Report for the year 1956-1957: Madhya Pradesh.

APPENDIX

TABLE 1—DATA FOR INDIAN AGRICULTURE^a

State	<i>T</i>	<i>K</i>	<i>L</i>	<i>w'</i>	Π	<i>V</i>
West Bengal	12.15	127.33	402.41	1.54	923.28	1,811.56
	16.96	116.00	628.37	1.61	772.36	2,403.23
	.64	7.44	39.05	1.60	187.78	129.41
	1.81	14.84	97.96	1.49	373.03	352.59
	3.11	25.19	173.10	1.59	555.87	547.05
	4.47	33.30	213.58	1.53	1,948.21	809.07
	6.18	41.59	321.42	1.45	813.20	1,158.13
	8.15	37.89	323.80	1.54	955.08	1,401.80
Madras	11.81	86.21	336.58	.54	1,653.61	907.01
	17.35	93.69	395.58	.56	2,215.54	1,174.59
	22.97	103.36	560.41	.62	2,248.45	1,683.70
	43.78	205.76	897.49	.55	5,838.73	3,607.47
	1.61	39.60	179.35	.62	426.00	354.04
	3.66	37.69	229.85	.52	716.90	751.03
	6.02	67.42	276.92	.56	2,045.88	947.55
	8.83	98.89	342.60	.56	763.14	1,190.28
Madhya Pradesh	12.44	9.57	294.70	1.08	1,709.28	1,479.12
	17.19	11.86	403.45	1.00	6,718.47	1,693.21
	24.25	14.55	470.21	1.11	40.53	2,616.57
	34.77	31.64	756.25	1.04	144.37	3,689.10
	45.17	41.10	1,084.08	1.11	157.86	4,458.28
	93.36	82.15	1,831.72	1.15	334.62	10,017.53
	2.95	3.42	101.13	.94	513.87	422.73
	7.38	8.63	190.40	1.06	729.34	849.44
Uttar Pradesh	12.00	78.00	602.40	1.06	7.57	2,448.00
	16.90	95.99	765.57	1.06	320.98	3,380.00
	27.58	148.93	1,073.14	1.01	384.68	5,653.90
	3.33	31.00	209.16	1.01	411.48	922.41
	7.68	64.97	432.84	.98	227.80	1,843.20
Punjab	14.50	19.57	450.22	1.51	448.19	2,463.55
	28.45	20.48	701.86	1.38	124.41	4,056.97
	81.19	30.85	1,484.96	1.92	391.14	12,957.92
	3.98	8.95	158.96	1.33	129.43	702.47
	7.45	7.37	270.88	1.40	377.94	1,270.22

^a For identification of variables, see Section IV of text.

The Incidence of Social Security Payroll Taxes

By JOHN A. BRITAIN*

One of the ironies of recent years is the sharp rise in the taxation of labor income in the midst of a declared war against poverty. As of 1970, a family of six earning a wage or salary of \$5,000 per year was officially classified by government agencies as living in a condition of poverty. Its plight is taken into account by the income tax system which completely spares such a family by means of exemptions and deductions. Although the social security system refers to its own levy as a "contribution," it is in fact an involuntary payroll tax without exemptions or deductions. A tax of \$240 is withheld from this \$5,000 family, and families farther down in the poverty range also pay 4.8 percent. Furthermore, the additional taxes for social security (including unemployment insurance) charged against employers average about 7 percent of covered earnings, and it is important to ask who is ultimately paying this portion of payroll taxes. If, as classical theory suggests, labor really bears both parts, the total payroll tax imposed on a family in the poverty range is about 12 percent—\$600 in the case of the \$5,000 family. Moreover, much steeper employer payroll taxes are imposed in many other countries (e.g., about 40–50 percent in

Italy), and the incidence of this type of tax is highly important in assessing labor costs to the employer and the tax burden on employees.¹

Although there is little empirical evidence on the burden of this tax, some economists believe that it is virtually axiomatic that the employer contribution is borne by labor in the long run. For example, Milton Friedman has written:

[The total tax for social security] includes what is euphemistically called "a contribution by the employer." Again, this is mislabeling. It is no contribution by the employer: it is a compulsory tax and it isn't paid by the employer; it is, in effect, paid by the wage earner. It is part of his wages that is sent to Washington instead of going to him. The form, the name, doesn't change the substance.²

Many other economists, social security specialists, employers, and labor unionists in this country appear to hold an agnostic view of the proposition that labor bears the employer tax. The tax is thus seen by some as mitigating the regressive impact of the tax on employees,³ and the Social Security Administration has been adamant in its belief that the employer tax should not be imputed to employees in comparing

* Senior Fellow, Brookings Institution. The issue considered here arose within a general study of the economic effects of payroll taxation under a program supervised by the National Committee on Government Finance, and financed by a special grant from the Ford Foundation. I am grateful for the suggestions of Joseph A. Pechman, Joan L. Turek, the referee, the managing editor, and for the research assistance of Sheau-eung Lau and Julia Clones. The interpretation and conclusions are my own and do not necessarily reflect the views of the staff, officers, or trustees of the Brookings Institution.

¹ The concepts of incidence and shifting adopted here stress the effects of the tax on aggregate factor shares. The degree of shifting measures the extent to which the effective incidence departs from the impact incidence (see Richard Musgrave, ch. 10).

² See Milton Friedman, p. 8.

³ The employee tax is generally regressive because the marginal tax rate is usually proportional below some earnings ceiling, and zero above it, and because the tax does not apply to property income which is concentrated in the higher brackets; a tax truly borne by employers would tend to offset this.

lifetime taxes and benefits under the social security system.⁴ Among economists, George Jaszi has concluded: "The shifting of social security taxes is a matter about which little is actually known" (p. 53). A well-known textbook by Harold M. Groves repeats this view in a forlorn summary of the state of knowledge:

The incidence of these taxes cannot be predicted with any great degree of assurance. On the whole, the safe conclusion would seem to be that it is divided among employers, labor and consumer; but in what proportion cannot definitely be said. [p. 157]

These conflicting opinions invite further investigation of the question, especially in view of the fact that the payroll tax has been the fastest growing major tax in the postwar United States. The maximum tax per employee for Old Age, Survivors, Disability and Hospital Insurance alone rose from \$60 in 1949 to \$749 in 1970 and is scheduled to reach \$1,017 in 1973. Expansion of the social security tax was also the primary factor in the rise of total "Contributions for Social Insurance" from \$5.7 billion in 1949 to about \$54 billion in 1969. The yield from this taxation of labor income substantially exceeded corporate income taxes in 1969, and the yield was about one-half that of all individual income taxes. "After hearing for years that social security benefits are too low. . . Congressmen are now deluged with complaints that social security taxes are becoming burdensome."⁵ The upward trend in payroll taxation is also spreading to the local level where city governments

⁴ "First, it is necessary to make clear that any such comparison should not include the employer contribution, because that contribution must necessarily be considered to be pooled for the benefit of all covered persons . . ." (U.S. House of Representatives, p. 331). This reason for excluding the payroll tax is unacceptable, since the distribution of *benefits* payable out of the receipts from the employer tax is irrelevant to the question of the incidence of the tax itself. It certainly does not justify treating the tax as a burden on no one.

⁵ *New York Times*, March 20, 1967, p. 15.

are showing increasing interest in payroll taxes as one means of snaring revenue from nonresidents employed in the city. Payroll taxation bulks far larger in Europe and many less industrialized countries, and Britain is now trying to move labor into manufacturing by means of a selective employment tax.

As this type of taxation grows, the unresolved question of the incidence of the employer portion (well over one-half of the U.S. payroll taxes) becomes increasingly important for a number of reasons. The location of the tax burden is relevant to appraisal of its effects on employment, income distribution, economic stability and growth. It is also essential to appropriate income tax treatment, appraisal of the rate of return on "contributions," rational collective bargaining, and the impact on migration and relative labor costs of competing economies.⁶

The present objective is to isolate quantitatively the impact of employer payroll taxes on factor shares in the long run. The primary approach builds on the relationship between the compensation and productivity of labor which has been frequently observed, and is also implied under certain assumptions by the specification of a production function. An inter-country analysis is undertaken which takes statistical advantage of the wide variation in tax rates among countries. The overall indication is that firms treat their payroll tax like any other labor cost in setting output and price and in agreeing to the total compensation to be awarded a given degree of labor productivity. The long-run result appears to be that employers in the aggregate avoid the burden of their contribution via a trade off between the tax and real wages and salaries.

⁶ A recent attempt to show that capital was not relatively lightly taxed in Western Europe rested entirely on an a priori assumption that payroll taxes are borne by capital. See Vito Tanzi, pp. 39-44.

I. Some Preliminary Conjecture and Casual Empiricism

Opinion and Practice in the Labor Market

Interpretation of their own behavior by participants on the economic scene is not notably reliable. However, it is worth asking how the parties in the labor market view the impact of the employer payroll tax. The issue could be expected to surface explicitly and frequently in these European countries in which the tax is much heavier than here. For example, a French union official clearly implied in a graphic opinion that the employer "contribution," in large part earmarked for family allowances, is in reality paid by labor: "We are getting paid less and less for our work, and more and more for being 'Father Rabbit,'" he complained.⁷ It is clear that he was taking for granted that the tax on employers was paid at the expense of the basic wage. A similar French union opinion was reported by Lorwin (1952, p. 362): "The real incidence of social and indirect wage charges made the working class function as a vast mutual aid association in which . . . it was the poor who were helping out the poorer." And as stated in James Vadakin (p. 131), "French employers frequently argue in collective bargaining that the imposition of social charges precludes their granting of wage increases." Even in Sweden where the tax is much smaller, it appears to have forestalled wage increases; for example, in 1959:

Swedish workers attached immense importance, during the wage talks to the pension plan which the Government was expected to adopt. . . . The unions felt, however, during the wage negotiations, that they could not get both a pension law and a substantial contractual wage increase this year, since the pension bill called for sizable employers' contributions to the pension fund, amounting to 1.89 percent of the payroll in 1960 and increasing to 4.2 percent in 1964.⁸

⁷ See Val Lorwin (1952) p. 362.

⁸ Bureau of Labor Statistics, p. 6.

In the United States explicit references to a trade off between the employer tax and private compensation are less common. However, direct recognition of the shifting of the tax to employees can be found, for example, in the academic world. At 35 percent of the colleges and universities affiliated with the Teachers Insurance and Annuity Association, the institution's contribution to this private pension fund is adjusted downward to offset the cost of each increase in the taxable ceiling under the social security program. Many collective bargaining agreements concerning private pension payments have included similar offsets; however, this practice is becoming rare. Even so, since employers are likely to assess their ability to pay in terms of profits after all such costs are netted out, tax increases cut expected profits and weaken labor's case for increases in the basic wage.

Some Casual Empirical Evidence

In an informative article, Majorie W. Hald found that International Labour Office data show an inverse relationship between the basic wage and the rate of "social charges" (employer payroll tax rates) in the manufacturing sector of thirteen European countries in 1954.⁹ This suggested that increased employer contributions might tend to be substituted for increases in the basic wage.¹⁰ Similar conclusions were reached by G. R. Reid (p. 111) on the basis of data for five countries, but this sample is so small that little can be read into the numbers. Other writers have looked at labor's real income

⁹ See Hald, pp. 33-35. No statistical test was provided, but the Spearman rank correlation coefficient is -0.50 , just significant at the 5 percent level on a one-tail test. (Only Ireland was far out of line, with a low wage level despite a low tax rate.)

¹⁰ This is, of course, a rather crude analysis, since other factors could have produced the observed statistical association. For example, governments in low-wage countries may be under great pressure to introduce public benefit programs. The more general model to be suggested here allows for such factors.

and relative share in the national income for clues on the incidence of the employer tax. Lorwin reports the following opinion on France in the early 1950's:

Despite understandable employer complaints about the burden of social charges, the increased benefits have come, not out of profits, but essentially out of a redistribution of income within the working class, as between direct wage and social wage recipients. . . . This phenomenon is made clearer by the fact that both (1) the total of *real* income going to wage earners (in the form of direct wages and social wage payments) and (2) wage earners' share in national income are about the same as in 1938. [1954, p. 46]

The stability of the ratio of labor's total compensation (including employer taxes) to national income, despite the great relative increase of employer taxes, is consistent with the proposition that the tax is substituted for the basic wage. However, this is a *ceteris paribus* interpretation, and other variables require consideration.

The above empirical reports are no more than suggestive. They are very persuasive however in comparison with the most recent empirical effort to appear,¹¹ which found that the share of property income in the national income of Puerto Rico fell with the introduction of the social security tax. This decline was deemed statistically significant on the basis of a misapplication of the chi-square procedure. This was then interpreted as showing that the employee tax (as well as the employer tax) was borne by capital even though the fall in the share of capital was four times too large to be explained by the tax alone. The reasoning and conclusions of that study are completely unconvincing.

II. Theoretical Analysis

Marginal Productivity Theory

The classical view that a universal employer tax is borne by labor was stated

succinctly by Harry G. Brown (pp. 160-63). He assumes that rationality and competition lead employers to hire workers until the point is reached where the wage is just barely recouped by the marginal value product. If a tax is imposed on the employer in proportion to labor hired, the marginal worker will no longer be hired unless he accepts a reduced wage which is lower by the amount of the tax. He can be expected to do so; his labor supply function is assumed to be highly inelastic, since he cannot hold out for long and has nowhere else to go under a universal tax. Brown denies that the tax could lead to higher prices, because he assumes the tax does not increase the money supply and would have little effect on aggregate demand.¹²

Brown's underlying assumption of a fixed amount of labor supplied is subject to question. The overall effect of a tax on labor supplied cannot be forecast with any confidence, since it depends on the unknown preference functions of individual workers with respect to income and leisure. The substitution effect tends to produce a contraction of labor supplied which may be counteracted by the income effect, but the relative strength of the two forces on balance is unknown. Certainly recent empirical studies in this area have produced nothing like a consensus, and this remains a qualification of Brown's case.¹³

¹¹ It is not clear why Brown rules out price increases supported by increases in aggregate money demand. If the tax proceeds are to be spent, an effort by employers to recoup the tax through price increases could be successful since the addition to government spending could yield enough total spending to buy the same output at the higher prices.

¹² Brown was considering a tax on labor with an assumed universal coverage—a structure approached fairly closely in this country. The incidence of a tax applied only to certain segments of the economy such as the British Selective Employment Tax concentrating on "services" is a more complex issue. Under certain simplifying assumptions such a tax, despite its allocative and price effects, may also be borne entirely by labor. For example, assume validity of the simple marginal productivity model, a world of two industries (one taxed, one not), two factors (labor and capital), unitary

¹¹ See Elizabeth Deran. For a gentle demolition of this article see Ronald Hoffman.

*Analysis Under Less
Restrictive Assumptions*

Theoretical analysis of the type outlined thus far has been criticized for over-simplified assumptions and dependence on the validity of the marginal productivity theory of wages.¹⁴ However, it is possible to restate the argument with greater generality and with less dependence on the various assumptions of the marginal productivity theory.

The theoretical case for imputing the employer tax to labor does not depend on profit maximizing behavior by employers. The less restrictive and more plausible assumption of cost minimization is sufficient, since it requires that the ratio of marginal value product to factor cost must be the same for all factors of production. With the imposition of an employer tax in relation to labor inputs, cost minimization cannot be maintained at the given level of employment unless the tax is recouped one way or another at the expense of the basic wage.¹⁵

It could be argued further, however, that the shifting of the tax to labor is not even dependent on the achievement of cost minimization, or competition in the labor market. Subject to qualifications considered later, it is only necessary to accept the concept of a demand curve for labor and the idea that it makes no difference to the employer what the sums (called "total compensation" in the national accounts) which he must pay to hire labor are called. The demand curve indicates that he will pay a certain price (measured

in real terms) for N units of labor of a given quality, and it makes no difference whether part of the price is called the employer's contribution and sent to Washington along with the employee's tax. There is no reason to expect a different employer reaction to the two components. Suppose the two taxes were suddenly combined into one package designated an employer tax and withheld by the employer as before. Except for short-run qualifications such as the current labor contract, there is no reason to expect the employer to pay more total compensation as the result of this accounting change. He would be required (in terms of formal accounting) to pick up the employee's part, and the nominal wage (total compensation minus employer tax) he would be willing to pay would be correspondingly reduced. A rational worker would be pleased to accept this cut in his nominal wage subject to federal and state income taxes, since the net result would be a slight increase in his after tax income. Because of the lower nominal earnings base the payroll tax and income tax paid out of the worker's given total compensation would be reduced.¹⁶ Furthermore, this argument does not depend on competition; it is equally true in noncompetitive labor markets where labor may receive less than its marginal value product.¹⁷

The purpose of the above argument is to support the one basic premise that the total real compensation which can be extracted for a given amount of labor (of given quality) is independent of the labels attached to the components. If this is accepted, and in addition the aggregate labor supply curve is completely inelastic, both payroll taxes are clearly borne by labor,

elasticity of demand for products, Cobb-Douglas production function, fixed supplies of factors and competitive conditions; under these assumptions, it can be shown that labor's share bears the exact amount of the tax despite allocative, price, and wage rate effects.

¹⁴ See Seymour Harris, pp. 291-99. See also other chapters in Part III of his work for detailed summaries and analysis of various incidence theories.

¹⁵ The complication of the picture by any employment effects is considered later.

¹⁶ There would be no corporate income tax effects, since wages and taxes are both deductible.

¹⁷ Even in a labor market with monopsonistic hiring and monopolistic labor supply, where the equilibrium is indeterminate, it does not seem plausible that the employer would distinguish among the components of compensation.

and there is no effect on the cost of labor or aggregate employment. The two taxes will be paid out of the total pre-tax compensation offered to labor at the fixed level of employment. Even if the supply of labor is elastic the same result (including no aggregate substitution of capital for labor) would occur if labor bargains in terms of total compensation, including both withheld "contributions," rather than after-tax income; that is labor views both taxes as part of its earnings, just as they are definitely part of the employer's costs. In this case the tax will still have no employment effect, and therefore there is no ambiguity about the outcome.

Only if two conditions hold simultaneously does an ambiguity arise. If the supply curve for labor were not perfectly inelastic *and* the supply price excluded the employer tax, the employer's attempt to recoup the tax from the basic wage would generally produce an employment effect. This result depends on labor's viewing one withheld tax as part of its income but not the other.¹⁸ This behavior, though difficult to rationalize, is possible, and the effect of the tax on labor's share would then depend on the elasticities of the labor supply and demand functions. Whether employment rose or fell, the basic wage bill could decline by more or less than the tax, depending on the change in the wage rate. Furthermore, we could not draw any particular conclusion such as a general tendency for labor to bear *less* than the full amount of the employer tax. This case should probably be considered a minor qualification of the general proposition, and even this minor indeterminacy depends upon an irrational view by labor that the employee tax is a price paid for future income

and services while the employer tax is paid for nothing. Note that the qualification is of less significance the more inelastic the supply of labor. The a priori case that the burden of both taxes is on labor seems very strong at the aggregate level.

Some Qualifications

Two types of qualification of the above reasoning should be indicated. First, the stress on real factor shares abstracts from the actual mechanism of the shifting process and makes no distinction between "backward" (money wage effect) and "forward" (price effect) shifting. While secondary to an analysis of the distribution of real incomes, the shifting process itself may affect the competitive position of economic units and regions. For example, if employers recoup the tax by price increases, exports would decline (unless exchange rates were adjusted).

A second qualification of the present reasoning is that its aggregative focus glosses over many likely microeconomic effects of actual payroll tax systems. Varying elasticities and rate differentials may produce a highly variable impact on different economic units and allocative effects of unknown importance. Still, although no light will be shed on these microeconomic effects of the tax, the question of its overall impact on labor's share seems of sufficient importance to warrant statistical investigation.

III. Hypotheses and Regression Models

Initial Hypotheses

The degree of shifting may be defined as the fraction s of the employer tax actually borne by labor. The a priori hypothesis is that the tax is borne entirely by labor; i.e., $s=1$. This follows from a competitive model of the demand for labor and an assumed zero elasticity of aggregate labor supply with respect to the real wage rate. Under pure "backward shifting" (no pro-

¹⁸ If the employee tax were also excluded from the supply price of labor, the same ambiguity would arise with respect to its incidence. The presumption that labor bears the employee tax only implies that the two taxes are not evaluated in the same way by labor.

duct price or employment effects from the tax), the fixed amount of labor would only be hired if the basic wage is lower by the amount of the tax than it would be in the absence of the tax. This is summarized in the first line of Table 1. If the assumption of zero supply elasticity is dropped *and* labor does not regard the employer contribution as part of its compensation, the resulting employment effect makes s dependent on the elasticities. In the special case of a unitary demand elasticity, the outcome would be determinate by definition. Since the aggregate compensation of labor (including the tax) would be invariant with respect to employment, the basic wage bill would be reduced by exactly the amount of the tax, and $s=1$ as indicated in the first column of Table 1. Thus even an employment effect does not necessarily rule out the proposition that labor bears the full burden.¹⁹

TABLE 1—DEPENDENCE OF HYPOTHESIZED VALUES OF THE SHIFTING COEFFICIENT s ON LABOR SUPPLY AND DEMAND ELASTICITIES

Elasticity of Labor Supply (Finite)	Elasticity of Labor Demand (Finite, Absolute Values)			
	1	>1	<1	0
0	$s=1$	$s=1$	$s=1$	—
>0	$s=1$	$s>1$	$0<s<1$	$s=0$
<0	$s=1$	$s<1$	$s>1$	$s=0$

The alternative hypotheses of $s>1$ and $s<1$ can be derived from plausible assumptions, as indicated in Table 1. Assuming a labor supply function with positive elasticity and a demand elasticity greater (less) than one, imposition of the tax would

reduce the basic wage bill by more (less) than the amount of the tax.²⁰ The more elastic the demand, the greater the negative employment effect relative to the positive compensation rate effect. In the case of a backward sloping labor supply curve, however, these relationships are reversed, since employment would increase (ignoring the unstable case in which the supply curve is less steep than the demand curve). With a demand elasticity greater (less) than one, imposition of the tax would reduce the basic wage bill by less (more) than the tax.

While values of s other than unity can be derived from plausible elasticities, the hypothesis $s=0$ (zero wage effect of the tax) seems highly implausible. This result depends on zero elasticity of demand with respect to the wage rate (in the case of an upward sloping supply curve). As summarized in Table 1, under any demand elasticity greater than zero, imposition of the tax would cut employment and the after-tax wage bill ($s>0$). J. R. Hicks has analyzed the determinants of the elasticity of derived factor demand at the industry level in a two-factor world (pp. 241–46 and pp. 373–78). Under the assumptions of the competitive marginal productivity model there is an exact relationship between the elasticity of the derived demand for labor e_d and a the elasticity of demand for the product, b the elasticity of the supply of capital, c the relative share of labor, and d the elasticity of substitution:

$$e_d = \frac{d(a+b) + cb(a-d)}{a+b-c(a-d)}$$

A zero elasticity of labor demand is implausible in itself, and this relation re-

¹⁹ Even this case is not entirely free of ambiguity concerning the impact of the tax, however. Assuming that firms were maximizing profits before the tax, the employment adjustment will only reduce—not eliminate the bite out of the share of capital. Thus the combined decline in the shares of labor and capital would be greater than the amount of the tax in this case of unitary demand elasticity.

²⁰ Since employment would be reduced by the tax, a demand elasticity greater than one would result (by definition) in a reduction in aggregate compensation (which includes the tax) and, therefore, a decline in the after-tax wage bill greater than the total employer tax ($s>1$).

inforces this a priori impression. At least two of the determinants of e_d (e.g., product demand elasticity and elasticity of substitution) must be zero in order for the expression to vanish. In view of the implausibility of such extreme values, the hypothesis $s=0$ seems untenable within the framework of the competitive model and an assumed upward sloping labor supply curve. Only if the latter is backward sloping is zero shifting possible with a demand elasticity other than zero, and this requires a special combination of elasticities. Even so, it seems in order to test $s=0$ as the logical alternative to $s=1$.

Cross-Section Regression Models

The objective here is isolation of the long-run impact of the employer tax on real wage rates, as distinct from the speed and process of shifting. Cross-section regression analysis of aggregative data for countries can offer direct evidence on the long-run response to the tax.²¹ The statistical models used were variations and elaborations of the estimating equation that emerges from the constant elasticity of substitution production function (CES).²² The variables originally considered in statistical analysis of the CES function were:

V = Value-added in thousands of U.S. dollars

L = Labor input in man-years

w = The basic wage rate (total labor cost excluding employer contributions for social insurance, divided by L), in dollars per man-year

The logarithmic transformation was gen-

erally favored, and the usual relationship estimated was:

$$(1) \quad \log V/L = a_1 + b_1 \log w + u_1$$

Assuming validity of the CES production function, competitive product and factor markets which are in equilibrium, correct measurement of variables and an exogenous wage rate, the estimate of the slope b_1 is an estimate of the elasticity of substitution between labor and all capital inputs. The claim that the estimate of b_1 is an estimate of this elasticity has been challenged on many grounds,²³ but this interpretation of the coefficient is not essential to the present application of the model. One can acknowledge the distinguished paternity of equation (1) as a point of departure without being dependent in any way on the rigid assumptions needed to deduce it from the underlying theoretical construct. The relationship is a commonsense one a priori, and indeed the original authors apparently tried it out statistically before coming up with the CES function as a theoretical underpinning. In addition, the particular direction of association specified in equation (1) does not follow from the underlying model, and the reverse specification appears equally plausible. To avoid dependence on any assumption on this score, the equations have been fitted both ways.

The assumed relationship was generalized to permit explicit analysis of the impact of the employer contributions for social insurance. The new specification can be most readily rationalized for the version which states (in logarithmic form) that the average wage rate in a country is dependent on the productivity of its labor:

²¹ For a statement of the argument that cross-section regressions do generally yield long-run relationships see, for example, Lawrence Klein, (1962) pp. 52-60. The stress on countries rather than states was dictated by institutional realities. There are enormous intercountry differences in employer payroll tax rates but only minute differences across the states of this country.

²² For the original presentation of the underlying theoretical construct, see Kenneth Arrow et al, pp. 225-50.

²³ For example, it has been argued that the CES function itself, like the Cobb-Douglas or models with fixed input coefficients, is only a special case of a more plausible and more general model. Recognition of this would require the presence of the capital variable in the estimating equation.

$$(2) \quad \log w = a_2 + b_2 \log V/L + u_2$$

The question immediately arises as to whether the wage rate associated with a given level of productivity includes the employer payroll tax per unit of labor in addition to the basic wage rate.²⁴ This element can be incorporated as an effective tax rate t applied to the basic (private) wage rate w . The rate of compensation of labor was assumed to be best measured by w plus some unknown fraction s of the employer payroll tax tw . This generalized model is:

$$(3) \quad \log w(1+st) = a_3 + b_3 \log V/L + u_3$$

For estimation of s , this may be rewritten with the basic wage rate as the dependent variable:

$$(3a) \quad \log w = a_3 + b_3 \log V/L - \log(1+st) + u_3$$

The coefficient s in this model may be interpreted as the "shifting coefficient," or the fraction of the employer tax per worker which is actually borne by labor. For example, if s should equal zero, $\log(1+st)$ equals zero, and the estimated basic wage rate would be independent of the tax, depending solely on productivity; this would indicate no shifting. However, an s value of unity would indicate that for a given level of productivity V/L the presence of the tax lowers the estimated basic wage rate by just the amount of the tax.²⁵ This shows a direct and complete trade off between the basic wage rate and the tax per worker, or a 100 percent shift-

ing of the tax burden at the expense of labor's basic wage.²⁶

The term $\log(1+st)$ is awkward for estimation purposes, but the parameter s can be extricated by approximating $\log(1+st)$ either by $s \log(1+t)$, or by the first term of the Taylor expansion $s(.434t)$;²⁷ results by the two methods agree closely. For estimation purposes two approximations of the last equation emerge:

$$(4) \quad \log w = a_4 + b_4 \log V/L - s_4 \log(1+t) + u_4$$

$$(5) \quad \log w = a_5 + b_5 \log V/L - s_5 (.434t) + u_5$$

Treatment of $\log V/L$ as the dependent variable leads by analogous reasoning to two alternative estimating equations in which s continues to be interpreted as the shifting coefficient:

$$(6) \quad \log V/L = a_6 + b_6 \log w + b_6 s_6 \log(1+t) + u_6$$

$$(7) \quad \log V/L = a_7 + b_7 \log w + b_7 s_7 (.434t) + u_7$$

These four models all rely on the logarithmic transformation of the variables. This form is preferred and stressed for methodological reasons.²⁸ However, the corresponding models without the transformation were also estimated in the early

²⁴ This shifting interpretation is, of course, subject to the qualification concerning possible employment effects discussed in Section II. The effect of the tax on the wage *bill* could be different in relative terms from its effect on the wage rate. However, any such complication signals no particular direction of bias in s as a measure of the degree of shifting.

²⁷ The first of these approximations is exactly correct when s equals zero or unity, too low for s values in between, and too high for s greater than one. The second is correct if $s=0$ and too high for positive values of s . The absolute error in the approximations varies positively with t but does not seem excessively large for extreme values of t such as 0.4. In any case the results from both approximations will be presented to suggest the order of magnitude of error.

²⁸ From the theoretical point of view, the logarithmic version might be favored because it is the one which emerges from the well-known CES rationale. It was also preferable for statistical reasons, since it afforded a closer approach to the property of homoscedasticity.

²⁴ Note that the basic wage rate w , as defined above, includes private fringe benefits and does not exclude payroll taxes paid by employees.

²⁵ If $\log w_0$ and $\log w_1$ are the regression estimates of the dependent variable before and after tax (and w_0 and w_1 are the implied basic wage rates), $\log w_1 = \log w_0 - \log(1+t)$, and $w_1 = w_0 + tw_1$. So, the new basic wage rate tends to equal the original basic wage rate less the tax per unit of labor. This interpretation may be generalized to cover values of s other than zero or one. The model implies that $w_1 = w_0 - stw_1$; the new wage rate tends to fall short of the old by the fraction s of the tax tw_1 .

stages, because they avoided the need to approximate $\log(1+st)$.²⁹ The equations are

$$(8) \quad w = a_8 + b_8 V/L - s_8 tw + u_8$$

$$(9) \quad V/L = a_9 + b_9 w + b_{9s} stw + u_9$$

Models (4)–(9) offer six alternative estimates of the shifting coefficient s .

IV. Empirical Findings

Most of the basic country data on value-added, wages and salaries, and employment were taken from the manufacturing censuses for 1958, but any annual census in the period 1957–59 was treated as eligible for the cross-section analysis. Three sets of countries and four sets of currency conversion ratios (called x_1 , x_2 , x_3 , and x_4) were used in fitting the models. Effective payroll tax rates were estimated from statutory rates. Data sources, criteria for country selection and methods of calculating the conversion ratios are outlined in the Appendix.

Table 2 reports estimates of the shifting coefficient s based on models (4)–(9) fitted to data for aggregate manufacturing.³⁰ All of the models produce close fits for all sets of data, with 92–96 percent of the variance explained in each case.³¹ The results for the logarithmic models (4)–(7) show posi-

²⁹ Equation (8) also offers the most direct interpretation of the coefficient s . For example, if the estimate of s turns out to be unity, the expression says that for a given level of productivity the higher the tax per unit of labor tw , the lower the estimated basic wage rate by the same amount.

³⁰ The fitting of models to aggregate data is clearly a distasteful procedure when viewed within the production function framework. However, the negative association revealed between the tax variable and the basic wage, given productivity, nevertheless seems meaningful and indicative of a trade off between basic wages and employer taxes.

³¹ Results are reported for conversion ratios x_1 and x_4 only, because they were very nearly duplicated by results for ratios x_2 and x_3 , respectively. It is also worth noting that the models using estimated purchasing power parity ratios x_1 and x_4 produced closer fits in every case than those for the relatively arbitrary official exchange rates contained in variables x_2 and x_3 .

tive values of s which are significantly greater than zero in every case (at the 2.5 percent level of significance or better). On the basis of this methodology, the no-shift hypothesis appears thoroughly discredited at the aggregate level. Although all 24 of the point estimates of s (falling in the range 1.14–1.60) are greater than the alternative hypothetical value of unity, they are not embarrassingly large; the estimates exceed one by a maximum of about one standard error and are therefore consistent with the hypothesis that 100 percent of the tax is borne by labor. The results for models (8) and (9) without the logarithmic transformation are somewhat more erratic, but tell the same story. These models show estimates of s significantly greater than zero in ten out of twelve cases. In only two cases were the estimates substantially greater than unity. In sum, models (8)–(9) support the shifting results of models (4)–(7) but less convincingly.

Wages and productivity in two-digit manufacturing industries were next analyzed for the maximum number of countries with data available; results for individual industries are presented in Table 3.³² Several of the industry models fit considerably less well than the aggregate models, but on the whole the \bar{R}^2 estimates remain very high (falling below 0.9 for only three industries). The estimates of the shifting coefficient s based on models (4)–(7) are significantly greater than zero at the 5 percent level or better in the case of 7 of the 13 industries, but in no case significantly greater than unity; thus the hypothesis that $s=0$ can be rejected in favor of $s=1$ is supported strongly in the majority of the industries. Only in the single case of the nonmetallic mineral products industry can the hypothesis that $s=1$ be rejected in

³² The purchasing power parity ratio x_1 produced closer fits than the exchange rate x_2 for all 13 industries. However, the estimates for s were so similar in the two cases that only those for x_1 are tabulated.

TABLE 2—REGRESSION ESTIMATES OF SHIFTING COEFFICIENTS BASED ON AGGREGATE DATA FOR ALL MANUFACTURING AND CONVERSION RATIOS x_1 AND x_2

Model	Statistic	64 Countries (Sets A, B, C)		44 Countries (Sets A, B)		30 Countries (Set A)	
		x_1	x_2	x_1	x_2	x_1	x_2
(4)	s	1.325 ^a	1.326	1.535 ^a	1.435 ^b	1.538 ^b	1.564 ^b
	$S(s)$	(.463)	(.467)	(.463)	(.486)	(.650)	(.691)
	\bar{R}^2	.930	.924	.934	.921	.944	.926
(5)	s	1.149 ^a	1.140 ^a	1.317 ^a	1.224 ^a	1.392 ^b	1.413 ^b
	$S(s)$	(.408)	(.411)	(.405)	(.425)	(.580)	(.616)
	\bar{R}^2	.929	.924	.933	.920	.944	.926
(6)*	s	1.517 ^a	1.533 ^a	1.597 ^a	1.471 ^a	1.527 ^b	1.561 ^b
	$S(s)$	(.473)	(.478)	(.478)	(.507)	(.673)	(.721)
	\bar{R}^2	.932	.926	.934	.920	.934	.925
(7)*	s	1.313 ^a	1.313 ^a	1.370 ^a	1.252 ^a	1.380 ^b	1.407 ^b
	$S(s)$	(.416)	(.421)	(.418)	(.444)	(.600)	(.644)
	\bar{R}^2	.931	.926	.933	.920	.943	.925
(8)	s	1.286 ^a	.892 ^a	1.484 ^a	1.004 ^a	1.046 ^a	.658
	$S(s)$	(.395)	(.454)	(.424)	(.530)	(.556)	(.678)
	\bar{R}^2	.943	.932	.948	.947	.960	.934
(9)*	s	1.706 ^a	1.428 ^a	1.762 ^a	1.418 ^b	1.310 ^b	1.083
	$S(s)$	(.386)	(.452)	(.414)	(.532)	(.550)	(.685)
	\bar{R}^2	.949	.938	.953	.933	.963	.937

Source of data: See Appendix.

* The coefficient s in models (6), (7), and (9) is estimated by the ratio of the coefficients bs and b . The standard errors $S(s)$ for these three models are derived from an abbreviated version of the Taylor's expansion approximation (see Klein (1953), p. 258). The estimate used was the ratio $S(bs)/b$. This approximation ignored the variance of b and the covariance of b and bs on the ground that they were generally small relative to the variance of bs .

^a Different from zero in the expected (positive) direction at the 5 percent level of significance.

^b Different from zero in the expected (positive) direction at the 2.5 percent level of significance.

^c Different from zero in the expected (positive) direction at the 0.5 percent level of significance.

favor of the hypothesis that $s=0$. The median estimate of s is that obtained for the clothing and footwear industry, whichever one of the four models is applied; though each of these estimates of the shifting coefficient s is slightly greater than one, none is significantly above. These industry results, while not unanimous, greatly favor the hypothesis $s=1$ over $s=0$ and indicate a 100 percent trade off between payroll taxes and the basic wage rate.

In an attempt to pin down the estimate for the aggregate shifting coefficient for manufacturing, the industry data underlying Table 3 were pooled for the final estimates of models (4) and (5). Dummy var-

iables were introduced to permit the constant term and coefficient of $\log V/L$ to vary by industry.³³ Results for the two models and two exchange conversion methods are presented in Table 4.

The dummy variable technique allows 380 degrees of freedom and should yield by far the most accurate estimates of s attained in this study, with standard errors reduced to about 0.25. The estimated values of the shifting coefficient in Table 4 are greater than zero by 4.0–4.5 standard

³³ Models (6) and (7) could not be treated in this way to yield a unique estimate of s because the estimate of s depended on the ratio of the single estimate of bs to the estimates of b which were assumed to vary by industry.

TABLE 3—ESTIMATES OF THE SHIFTING COEFFICIENT s BASED ON
INDUSTRY DATA AND CONVERSION RATIO α_1

Industry	Number of Countries	Statistics	Model (4)	Model (5)	Model (6)	Model (7)
Food, Beverages and Tobacco	36	s	1.74 ^a	1.57 ^a	1.83 ^a	1.65 ^a
		$S(s)$	(.86)	(.78)	(.90)	(.81)
		\bar{R}^2	.907	.907	.907	.907
Textiles	34	s	1.82 ^b	1.64 ^b	1.99 ^b	1.78 ^b
		$S(s)$	(.86)	(.77)	(.89)	(.79)
		\bar{R}^2	.929	.929	.930	.930
Clothing, Footwear, etc.	31	s	1.23 ^b	1.12 ^b	1.23 ^b	1.12 ^b
		$S(s)$	(.52)	(.46)	(.53)	(.47)
		\bar{R}^2	.971	.972	.971	.971
Wood and Products	36	s	1.64 ^b	1.52 ^b	1.82 ^b	1.67 ^b
		$S(s)$	(.79)	(.71)	(.81)	(.72)
		\bar{R}^2	.933	.933	.934	.934
Pulp and Paper Products	28	s	.77	.71	.73	.67
		$S(s)$	(1.00)	(.89)	(1.04)	(.93)
		\bar{R}^2	.918	.918	.917	.918
Printing and Publications	34	s	1.37 ^a	1.20 ^a	1.43 ^b	1.25 ^a
		$S(s)$	(.67)	(.60)	(.69)	(.61)
		\bar{R}^2	.946	.946	.946	.946
Leather and Products	25	s	-.16	-.08	-.34	-.23
		$S(s)$	(.79)	(.70)	(.82)	(.72)
		\bar{R}^2	.929	.929	.929	.929
Rubber Products	27	s	1.15	1.07	1.13	1.05
		$S(s)$	(1.36)	(1.20)	(1.46)	(1.29)
		\bar{R}^2	.859	.859	.858	.859
Chemicals and Products	31	s	.89	.80	1.16	1.01
		$S(s)$	(1.60)	(1.43)	(1.80)	(1.61)
		\bar{R}^2	.771	.771	.772	.772
Nonmetallic Mineral Products	37	s	-.33	-.24	-.32	-.23
		$S(s)$	(.71)	(.64)	(.73)	(.66)
		\bar{R}^2	.942	.942	.942	.942
Basic Metals	25	s	2.07 ^a	1.86 ^a	2.20 ^a	1.97 ^a
		$S(s)$	(1.14)	(1.02)	(1.22)	(1.08)
		\bar{R}^2	.879	.879	.879	.879
Metal Products	35	s	1.56 ^b	1.40 ^b	1.67 ^b	1.48 ^b
		$S(s)$	(.69)	(.62)	(.70)	(.63)
		\bar{R}^2	.943	.943	.944	.944
Other Manufacturing	28	s	1.02	.92	1.04	.93
		$S(s)$	(1.17)	(1.04)	(1.22)	(1.09)
		\bar{R}^2	.907	.908	.907	.907

Source of data: See Appendix.

^{a,b} For levels of significance, see Table 2.

errors, permitting the hypothesis that $s=0$ to be rejected at the 0.003 percent level or better. Again, although each estimate falls slightly above unity, the excess is far from significant, and the results strongly support the hypothesis that in the aggregate the entire employer tax is shifted to labor.^{34,35}

³⁴ It is possible that the s coefficients contain a slight

V. Some Implications of the Findings

It should be reiterated at this point that this analysis has shed no light on the mech-

upward bias for another reason. Insofar as other indirect taxes, such as the sales tax, bear more heavily on labor income than on overall value-added at factor cost and such taxes are correlated with payroll taxes across countries, the regressions on the payroll tax variable alone could pick up some of the influence of these other taxes; this would impart an upward bias to the estimates of s . However, the payroll tax is undoubtedly the domi-

TABLE 4—ESTIMATES OF THE SHIFTING COEFFICIENT s
FOR ALL MANUFACTURING ON POOLED INDUSTRY
DATA, WITH DUMMY VARIABLES
FOR INDUSTRIES

(407 Observations and 380 Degrees of Freedom)

Conversion Ratio	Statistic	Model (4) Plus Dummies	Model (5) Plus Dummies
α_1	s	1.144	1.043
	$S(s)$	(.263)	(.235)
	\bar{R}^2	.911	.911
α_2	s	1.176	1.070
	$S(s)$	(.268)	(.239)
	\bar{R}^2	.904	.904

Source of data: See Appendix.

anism through which the real burden of the tax on employers is shifted to employees. Earnings and productivity variables for each country were measured in dollars by several conversion methods. The essence of the finding here is that given the level of productivity in a country, the presence of a payroll tax on employers tends to reduce the wage in dollars by roughly the amount of the tax. This could be due to a lag in the basic wage (measured in local currency) in response to imposition of the tax (backward shifting); it could be due to price increases reducing the real value of wages (forward shifting). More likely, the outcome is achieved through a combination of the alternative employer reactions. The nature of the blend is of little significance for a study of income distribution, but the extent of price adjustment does affect international competitive positions until offset by exchange rate adjustment. In any case, whichever shifting mechanism dominates, the real burden of the tax falls on labor; this has important implications however it comes about.

nant tax concentrating on labor, and the impact of other taxes should be relatively small.

³⁸ Various forms of regression models allowing for lagged response were applied to U.S. time-series data. The findings gave considerable support to the cross-country conclusions but were not conclusive enough to merit reporting here.

If the conclusion that both employer and employee payroll taxes are borne by labor is accepted, several significant corollary propositions follow. In the first place, assuming no aggregate employment effects, this suggests that payroll taxes are neutral with respect to the allocation of capital and labor in the aggregate and within a given industry.³⁸ If the tax has no net cost impact for employers, it produces no incentive to substitute capital for labor. The conclusion that its burden falls on labor shields the tax from the usual criticism that it promotes automation and aggravates the unemployment problem.

The economic stabilization properties of the tax are also affected by its incidence. Although it is generally conceded that the typical payroll tax is a relatively weak stabilizer, it would probably be even weaker if borne by capital. For example, assuming a higher marginal propensity to spend income from labor than income from capital, a fall in the tax on labor in response to a wage decline would be a more effective brake to limit the decline in spending than an equal decline in a tax on profits. A tax borne by labor would also presumably produce a lesser drag on growth than a tax on profits, which would cause a greater cut in saving.

Acceptance of the labor burden hypothesis is also relevant in the collective bargaining arena. It should be recognized on both sides that the employer payroll tax is just as clearly a component of the cost of hiring labor as private fringe benefits or the nominal wage itself. Labor would then regard the employer contribution as part of its compensation which is being paid in lieu of a higher nominal wage. Recognition of a trade off between wages and fringes on the one hand and employer contribu-

³⁸ A payroll tax which varies across industries (such as the British Selective Employment Tax) can be expected to affect the allocation of labor among industries without reducing the overall capital-labor ratio.

tions on the other would bring into more explicit focus the pros and cons of fueling social programs by this type of tax.

The appropriate treatment of labor income under the personal income tax also depends on payroll tax incidence. At present the employee pays tax on the income from which employee contributions are withheld but not on the income from which the employer tax is withheld. If the latter income is part of labor's share it should be taxed just as the source of the employee tax is; or both parts should be exempt in favor of a tax on benefits which are now exempt.

If labor ultimately pays the employer tax, this is also highly relevant to relative international competitive positions. Countries such as Italy and France with large "social charges" are not placed at a competitive disadvantage vis-à-vis countries where the employer tax constitutes only a small part of total compensation.³⁷ This has significant implications for tariff policy and for attempts to improve international economic cooperation.

On the previous counts the labor-burden finding implies no significant critique of the payroll tax. However, the incidence of the tax is highly significant for evaluation of its effect on income distribution. The conclusion that labor bears the tax makes clear that its burden on low income groups is greater than generally realized. It also implies that its impact on income distribution is typically regressive. These qualities of the payroll tax offer a solid basis for proposing that this form of taxation be curtailed or eliminated. This could be done by introducing exemptions of low incomes as under the income tax or substituting the

income tax for all or part of the payroll tax.

Finally the incidence of the tax is significant for evaluation of the terms of social security programs. The finding that labor bears the tax points to a lower rate of return on contributions to participants in social security than if the employer tax could be ignored. It is difficult to understand the position of the Social Security Administration which has conceded that this tax is largely borne by labor in the aggregate and yet ignores it in evaluating the tax paid by individuals. It does so on the ground that no exact imputation of the tax is possible. However, if it is paid by employees as a group, it must also be paid by them as individuals, and it seems better to make imperfect imputations which are roughly right than to settle for being precisely wrong. The implication of these imputations is that wage and salary earners pay the entire tax for Unemployment Insurance and twice as much under the OASDHI program as the amount withheld from their nominal earnings. An awareness of this on the part of taxpayers might contribute to decreased reliance on this regressive form of taxation.

APPENDIX

Virtually all of the census of manufactures data on wages and salaries, value-added, and employment were taken from the United Nations publication, *The Growth of World Industry, 1938-61*. For a few countries, information was taken from other United Nations publications, *The Statistical Yearbook*, and *The Yearbook of National Account Statistics*.

The effective employer tax rate t was estimated from statutory rates in the Social Security publication *Social Security Programs Throughout the World*. Five types of employer payroll taxes were included; they were ear-marked for: (1) old-age, invalidity and survivors insurance and related programs, (2) health and maternity insurance,

³⁷ Even if the shifting of the burden to labor were accomplished primarily via price increases, this should have only temporary effects. The essential fact remains that the tax does not increase the real cost of labor; exchange rate adjustments could restore the competitive position which existed before any forward shifting.

(3) unemployment insurance, (4) family allowance programs, and (5) work injuries insurance. Estimates of effective rates took account of the taxable ceiling in each country. The statutory rate was adjusted on the basis of a graphical relationship between percentage of earnings taxable and the ratio of the ceiling to mean earnings, as observed in the United States. In the case of those countries that also specified minimum taxable income each earner was assumed to earn at least the minimum, and the effective rate was adjusted downward on the basis of the fraction of the total wage bill that was exempt by the minimum.

Four alternative sets of conversion ratios were used— x_1 , x_2 , x_3 , and x_4 —to convert currencies into dollars. The estimates x_1 and x_4 were "purchasing power parity ratios" based on price indexes; they were estimated by the United Nations, see *The Growth of World Industry, National Tables*, pp. 310–11, and *International Analysis and Tables*, Table 9B. These conversion ratios produced generally closer fits than the more arbitrary sets of official exchange rates x_2 and x_3 . These two sets (which differ slightly due to alternative treatment of multiple exchange rates) were based on U.S. Department of Commerce data and the U.N. *Statistical Yearbook*, Table 9A.

Countries were selected for analysis if data on value-added V , wages and salaries W , and employment L were available for aggregate manufacturing in at least one of the years 1957–59; most of the censuses were for the year 1958. A few additional countries were included for which the census fell just outside the 1957–59 period and for which the "number engaged" was available rather than the number of employees. This yielded data on aggregate manufacturing for sixty-four countries labeled sets A , B , C in the text. After fitting the models to these 64 observations, the 20 countries with the smallest total wage bills, set C , were dropped and the models refitted. Finally countries in set B with data on number engaged only, with surveys outside the 1957–59 period or without data available in the main source were dropped, leaving 30 observations. This process was an attempt to utilize successively

more reliable data while sacrificing observations. However, results for the three different sets of aggregative data all gave results consistent with the hypothesis that the overall shifting coefficient equals unity, as shown in Table 2.

The industry analysis in Tables 3 and 4 covers all countries among the original sixty-four for which data were available on an industry basis.

REFERENCES

- K. Arrow, H. B. Chenery, B. S. Minhas, R. M. Solow, "Capital-Labor Substitution and Economic Efficiency," *Rev. Econ. Statist.*, Aug. 1961, 43, 225–50.
- H. G. Brown, *The Economics of Taxation*, New York 1924.
- E. Deran, "Changes in Factor Income Shares Under the Social Security Tax," *Rev. Econ. Statist.*, Nov. 1967, 49, 627–30.
- M. Friedman, "Transfer Payments and the Social Security System," *The Conference Board Record*, New York, Sept. 1965, 11, 7–10.
- H. M. Groves, *Financing Government*, 6th ed., New York 1965.
- M. W. Hald, "Social Charges in the EEC Countries: Some Economic Aspects," *Econ. Int.*, Nov. 1959, 12, 677–96.
- S. Harris, *Economics of Social Security*, New York 1941.
- J. R. Hicks, *Theory of Wages*, 2nd ed., New York 1963.
- R. F. Hoffman, "Factor Shares and the Payroll Tax: A Comment," *Rev. Econ. Statist.*, Nov. 1968, 50, 506–08.
- G. Jaszi, "A Critique of the United States Income and Product Accounts," *Nat. Bur. Econ. Res. Stud. in Income and Wealth*, Vol. 22, Princeton 1958, p. 402.
- L. Klein, *A Textbook of Econometrics*, New York 1953.
- , *Introduction to Econometrics*, New York 1962.
- R. Lester, *Economics of Unemployment Compensation*, Princeton 1962.
- V. R. Lorwin, "France: History of Trade Union Developments," in W. Galenson, ed., *Comparative Labor Movements*, New York 1952, 313–409.

- , *The French Labor Movement*, Cambridge 1954.
- R. Musgrave, *The Theory of Public Finance*, New York 1959.
- G. R. Reid, "Supplementary Labour Costs in Europe and Britain," in G. R. Reid and D. J. Robertson, eds., *Fringe Benefits, Labour Costs and Social Security*, London 1964.
- V. Tanzi, "Tax Systems and Balance of Payments: An Alternative Analysis," *Nat. Tax. J.*, Mar. 1967, 20, 39-44.
- J. C. Vadakin, *Family Allowances*, Miami 1958.
- International Labour Office, *The Cost of Social Security, 1958/1960*, Geneva 1964.
- New York Times*, Mar. 20, 1967.
- United Nations, *Growth of World Industry, 1938-61, International Analysis and Tables*, New York 1963.
- , *Statistical Yearbook 1965*, New York 1967.
- , *The Growth of World Industry 1938-61, National Tables*, New York 1963.
- , *Yearbook of National Account Statistics, 1963*, New York 1965.
- U. S. Bureau of Labor Statistics, *Labor Development Abroad*, Washington, August 1959.
- U.S. Department of Commerce, *Statistical Abstract of the United States*, Washington 1958, 1959, and 1960.
- U.S. Department of Health, Education, and Welfare, *Social Security Programs Throughout the World, 1958*, Washington.
- U.S. House of Representatives, Hearings before the Committee on Ways and Means, *President's Proposals for Revision in the Social Security System*, Mar. 1, 2, and 3, 1967, Washington.

Determinants of the Commodity Structure of U.S. Trade

By ROBERT E. BALDWIN*

Nearly twenty years ago Wassily Leontief made the surprising discovery that a lower capital-labor ratio was required to produce a representative bundle of *U.S.* exports than was involved in producing a representative bundle of import-competing goods. The Leontief results and those from similar investigations of other countries¹ effectively destroyed the comfortable confidence of economists in the simple version of the Heckscher-Ohlin trade theory that had long been accepted mainly on the basis of casual empiricism.² However, the "Leontief paradox" also stimulated extensive theoretical and empirical research directed at providing alternative explanations for the commodity-pattern of a country's trade. The purpose of this paper is to test the main alternative hypotheses that have been advanced for this purpose, as well as the simple Heckscher-Ohlin theory itself, by using 1958

U.S. labor, capital, and input-output coefficients rather than the 1947 coefficients employed by Leontief. Information from the *1/1000 Sample of the Population of the United States, 1960* plus certain other data relating to the quality of the labor force and various structural characteristics of *U.S.* industries are also utilized in testing these hypotheses. In addition, the *U.S.* trade pattern for 1962 rather than 1947 or 1951 is used in making the various calculations.

The Heckscher-Ohlin theorem states that a country's exports use intensively the country's relatively abundant factors. As is well known (see the survey article by Jagdish Bhagwati 1965), a set of sufficient conditions for the theorem are: 1) identical production functions throughout the world for each commodity as well as qualitatively identical productive factors; 2) production functions homogeneous of degree one with diminishing marginal productivity for each factor; 3) nonreversibility of factor intensities; 4) identity of consumption patterns (in the sense that all goods are consumed in the same proportions) among countries at any given set of international commodity prices; and 5) perfect markets, free trade, no transport costs, and complete international immobility of productive factors. If one adds the condition that there are at least as many commodities as factors and that all countries produce some of each commodity, it also follows that factor-price equalization is achieved.³ In addition, as

* Professor of economics, University of Wisconsin, Madison. Research for the paper was financed by a grant from the National Science Foundation.

¹ See Masahiro Tatamoto and Shinichi Ichimura, Donald F. Wahl, Ranganath Bharadwaj, and Karl W. Roskamp.

² The simple Heckscher-Ohlin theory is a model in which only capital, labor, and natural resources are the factors of production, and in which such factors as economies of scale and differences in technology do not play a part in determining comparative-cost differences among nations. It should be recognized, however, that Bertil Ohlin, even when presenting what he regarded as a simplified version of his model, divided labor into three skill groups and capital into long-term and short-term capital. Nevertheless, over time most economists have come to label trade models with a two- or three-factor breakdown as simplified versions of the Heckscher-Ohlin theory. More important, they have believed that the broad pattern of a country's trade could be adequately explained with a simple two- or three-factor model. See, for example, Karl-Erik Hansson.

³ Factor-price equalization can, of course, be achieved without the identical-taste assumption. Moreover, it is

Pranab Bardhan has pointed out, the Heckscher-Ohlin theorem holds as long as one country's production functions all differ from those used in the rest of the world by only a neutral efficiency factor.

Obviously, the various conditions required for the Heckscher-Ohlin proposition to be logically valid do not all hold in the real world. However, this does not necessarily mean that the Heckscher-Ohlin theory is a poor theory. If the failure of any of the assumptions to hold does not systematically and significantly bias the conclusions of the model, the theory will generally still accurately predict the nature of trade patterns from a knowledge of relative factor endowments and thus be a "good" theory. Tests such as the one undertaken by Leontief are designed to determine the adequacy of the theory in this sense.

I. Alternative Explanations of the Leontief Paradox

Fortunately, trade theory has not suffered from a lack of suggested explanations for the Leontief results. Instead, the problem has been to discriminate among the several hypotheses that have been advanced to account for them. Six major (not necessarily mutually exclusive) groups of explanations can be distinguished.⁴ These maintain that the actual structure of *U.S.* trade can be accounted for mainly by: 1) the relative abundance of skilled labor in the United States; 2) an efficiency advantage in favor of the United States in Research and Development (R and D) oriented industries; 3) the scarcity of natural resources in the United States coupled with a complementary relationship between natural resources and capital; 4) factor-intensity reversals suffi-

ciently extensive to upset the Heckscher-Ohlin proposition; 5) a strong *U.S.* demand bias in favor of capital-intensive goods so that these are imported even though the United States is capital-abundant; 6) high tariffs and other trade-distorting measures that favor the domestic production of labor-intensive products and consequently bias the import bundle against these products.

Skilled Labor

Current interest in the general topic of investment in human resources has served to focus considerable attention recently (see articles by Peter Kenen (1968), Bharadwaj and Bhagwati, and Roskamp and Gordon McMeekin), on the first explanation mentioned above. This explanation was initially put forth by Leontief and Irving Kravis, both of whom pointed out that *U.S.* export industries employed more highly skilled labor than did import-competing industries. Donald Keasing (1965, 1966, 1968), Kenen (1965), Helen Wachrer, and Merle Yahr have since elaborated both analytically and empirically upon the significance of differential supplies of labor-skills and also demonstrated the importance of this factor for explaining trade patterns of other countries. Kenen has performed the interesting experiment on *U.S.* data of capitalizing the excess of wages earned by various types of skilled labor above the wages of unskilled laborers in order to obtain an estimate of the value of human capital involved in export- and import-competing production. When the estimates of human capital obtained by discounting at less than 12.7 percent are added to Leontief's physical capital estimates, the paradox is reversed.⁵

A drawback of computing human capi-

not necessary for factor-price equalization that each country produce some of every commodity.

⁴ See Gary C. Hufbauer for a classification that further refines some of the categories listed.

⁵ However, similar estimates by Bharadwaj and Bhagwati for India have the effect of operating against what would be predicted by the Heckscher-Ohlin theory.

tal by capitalizing income differentials at a single discount rate is, as Kenen notes, that the method assumes all income differences to be the result of differences in education and other forms of human investment. It also assumes that long-run equilibrium conditions prevail in capital markets.⁶ There is considerable evidence that market imperfections due to various economic and social factors as well as differences in ability are significant explanatory variables for earning differentials. Moreover, returns to low levels of education are considerably greater than to high levels of education.⁷

An even more important point is whether it is proper to combine estimates of human and physical capital to determine the capital-labor ratio in trade-oriented production. Such a procedure rests upon the assumption that in the long run, capital moves freely between physical goods and human agents of production. This assumption may be acceptable for a highly developed country like the United States, but it does not seem appropriate for most developing nations where market imperfections even make it difficult to regard all physical capital as fungible in the long run.

R and D Oriented Industries

Keesing (1968) together with Raymond Vernon, William Gruber, and others, has pointed to the significance of research activities in explaining trade patterns. In particular, they found that there is a strong positive correlation between the relative importance of R&D activities in American industries and the exports of American industries as a proportion of

total exports of all the major trading countries. These results confirmed the hypothesis that R&D expenditures are a proxy for temporary, comparative-cost advantages provided by the development of new products and productive methods.⁸ A more direct method of introducing non-uniform, efficiency differences between domestic and foreign production functions has been followed by Gary Bickel in a study of U.S.-Japanese trade. Bickel used differences among U.S. and Japanese industries in the productive efficiency of capital and labor (derived from empirically estimated CES production functions) as an explanatory variable for differences between the two countries in relative commodity prices. He found that at least 25 percent of the total variation in these prices was attributable to the efficiency factor alone.⁹

Scarcity of U.S. Natural Resources

The third-factor (natural resources) explanation of Leontief's results has been put forth both by Muhammad Diab and by Jaroslav Vanek (1963). Vanek accepts the notion that the United States is capital abundant but states that natural resources and capital are complementary. Therefore, since natural resources are scarce in the

⁸ This hypothesis also assumes that current R&D expenditures are representative of the stock of innovations that are the source of comparative-cost advantages and that the rate at which innovations are copied is approximately the same among industries.

⁹ Leontief's explanation (1953, p. 344) of his findings, namely that U.S. labor is some three times more efficient than foreign labor, implies that the efficiency advantage of the United States is highly biased towards saving labor whereas the usual intercountry studies that estimate the elasticity of substitution explicitly assume only factor-neutral efficiency differences. Bickel, for example, uses the neutral efficiency parameters calculated by Kenneth Arrow, Hollis Chenery, Bagicha Minhas, and Robert Solow. There is evidence suggesting that technical progress in the United States has in fact been labor-saving. See Paul David and Th. van de Klundert. However, the very large advantage in favor of U.S. labor suggested by Leontief does not seem to be supported by direct studies of comparative labor efficiency. See the article by Mordechai Kreinin.

⁶ By excluding the human capital in laborers, Kenen's method understates the total human capital involved in export- and import-competing production. Since laborers are more important in import-competing than export production, this exclusion has the effect of tending to reverse the paradox.

⁷ See Gary Becker (pp. 124-27) and Giora Hanoch.

United States, both capital and natural resources are conserved through trade. This hypothesis seemed to receive support from calculations that Leontief made in his second article (1956) on the subject. Specifically, when nineteen natural resource industries were excluded from the matrix, the paradox was eliminated.¹⁰

As William Travis has pointed out (pp. 94-99), this explanation is logically inconsistent with a Heckscher-Ohlin model in which factor-price equalization is achieved.¹¹ Even though there are other general or specific factors besides labor and capital, a country that is capital abundant relative to the rest of the world will export capital-intensive products compared to its import-competing production. This can be seen in the following way. In a factor-price equalization model where all goods are traded and tastes are identical, the factor proportions used to produce any particular commodity are the same in all countries, and each country consumes all commodities in the same proportions. This implies that each country indirectly consumes each factor in the same proportions. In other words, one can think of each country as starting with given factor supplies and then trading these at common factor prices until a common set of factor-consumption ratios is reached. In the case where natural resource industries are capital-intensive and natural resources are scarce in a capital-abundant country, factor equilibrium may be achieved in two ways: by the capital-abundant country

exporting items that are even more capital-using than natural resource products; or by the rest of the world, which is labor-abundant, exporting highly labor-intensive commodities in addition to the capital-intensive, natural resource products.¹²

Reversals of Factor Intensity

One of the most potentially damaging arguments against the Heckscher-Ohlin theory is that factor-intensity crossovers are extensive within relevant ranges of factor prices. Under these circumstances a country's exports to the rest of the world and the rest of the world's exports to the country may be either both capital-intensive or both labor-intensive. Then the Heckscher-Ohlin relationship cannot possibly hold for both trading units. A study by Minhas seemed to indicate that factor-intensity crossovers were in fact extensive. However, subsequent analysis by Leontief (1964), using additional data provided by Minhas, found extremely little evidence of factor-reversals. Several other recent studies¹³ also failed to support the Minhas position, but the matter cannot as yet be regarded as finally settled.¹⁴ For example, see the issues raised by Michael Hodd.

¹² Suppose that country *A* possesses a relatively abundant supply of capital, *K*, compared to country *B* (the rest of the world) in terms of either labor, *L*, or natural resources, *NR*, i.e., $K_a/L_a > K_b/L_b$ and $K_a/NR_a > K_b/NR_b$. Because of the equilibrium conditions that the consumption ratios of these factors must be equal and the value of a country's indirect exports of factors must equal its indirect imports of factors, it follows that country *A* must in effect export capital to country *B*. In a factor-price equalization model, this in turn implies that country *A*'s exports will be capital-intensive compared to its imports.

¹³ See the book by Hal Lary and the articles by Gordon Philpot and Merle Yahr.

¹⁴ The possible lack of global univalence between factor prices and commodity prices when goods are produced with specific natural resources as well as with capital and labor should be further investigated. This corresponds to factor-intensity reversals in the two factor-two commodity case. James Ford (p. 60) raises this point.

¹⁰ The industries excluded by Leontief tend to be those in which the direct and indirect factor content of immobile natural resources is relatively high. His exclusion of all agricultural industries except livestock and livestock products seems questionable under this criterion.

¹¹ Vanek (1968) also has now proved this proposition rigorously. Travis (p. 97) further notes that the Leontief results are not consistent with the natural resource explanation even in the absence of factor-price equalization.

Demand Bias

Demand bias is invariably cited as a possible explanation of the Leontief paradox but no writer has strongly argued that this is the major explanation. Indeed it is now usually accepted (see, for example, the article by Arthur Brown and the analysis by Travis, pp. 105-10) that, if final demand in the United States is factor-biased, the bias is towards labor-intensive rather than capital-intensive goods, because of the operation of Engel's law.

Tariffs and other Distortions

The argument that various tariff and nontariff trade-distorting measures account for Leontief's results has been expounded most cogently by Travis. He arrives at this conclusion after carefully showing the correctness of Leontief's test of the Heckscher-Ohlin theory and then arguing that it is highly unlikely that the failure of the various assumptions of the Heckscher-Ohlin theory to hold (other than the free trade one) could account for Leontief's results.¹⁵

Another explanation stressing the importance of market imperfections is Diab's suggestion (pp. 53-56) that commodities produced abroad by American corporations or their subsidiaries and with the aid of American capital, know-how, and highly skilled technicians and managers should be regarded as part of U.S. internal trade rather than imports. Since a large part of this production consists of capital-intensive, natural resource products (especially minerals), the paradox might well be reversed if these were excluded from the trade pattern. However, as Travis points out (pp. 110-11), for this argument to

be valid it is necessary to explain why American capital, once overseas, does not move into labor-intensive industries in foreign countries.

Although Travis seems to believe this point cannot be explained satisfactorily, there is considerable evidence in the literature on economic development supporting the view that the immobility of foreign capital and know-how between the export and domestic sectors of less developed countries is a real phenomenon and is based on economic factors. Consider, for example, why foreign funds have in the past flowed mainly into export-oriented, natural resource industries in the less developed countries or in tertiary lines that serve to support these industries. Part of the explanation is that there usually is better knowledge in the developed countries concerning profit opportunities in the developing countries with regard to natural resource industries compared to most other products. Because of the generally lower supply elasticities in developed countries for natural resource products than for commodities produced mainly with capital and labor, there tends to be a greater upward pressure on the prices of those natural resource products that are significant inputs into industrial processes than on the prices of most other products, as growth takes place in the advanced countries. This relative price movement alerts investors to the obvious profit opportunities that can be exploited if costs can be kept from rising and thereby leads to a search at home and abroad for new supply-sources as well as for better ways of using existing supply sources. On the other hand, even if highly profitable opportunities exist overseas in product lines outside of the natural resource group, investors are less likely to become aware of them because of the absence of this signaling mechanism.

Other important factors affecting for-

¹⁵ Tariffs can weaken the pattern of indirect factor-trade in a Heckscher-Ohlin model but cannot alone produce paradoxical results. Export subsidies (or some domestic distortion) in lines that intensively use a country's relatively scarce factors are needed to produce these results.

eigners' decisions to invest in less developed countries are the nature of factor supplies in these countries, the size of markets, and the degree of input-complexity of production. Natural resource conditions are often sufficiently favorable to make foreign investment profitable in large scale, primary industries that can supply the large markets of developed countries. However, for products that do not rely heavily on the natural resource factor, production costs usually are too high for exports to be internationally competitive. Labor with very little skill is abundant, but without some training this labor is very inefficient even when used in producing the simplest types of manufactures under modern methods.

A lack of trained labor is less of a barrier to the competitive production of manufactured goods for domestic consumption. However, the costs of establishing and supervising productive units abroad tends to be prohibitive unless the optimum plant size is large. But, in industries where the optimum size of productive units is large, domestic demand usually is too small to support efficient production. Still another factor discouraging foreign investment is the more complex system of input requirements (direct and indirect) for manufactures than for primary products. It is more difficult to finance, coordinate, and fully utilize interdependent investment projects in several as compared to a few industries.

After foreign capital moves into export-oriented, natural resource industries in less developed countries, it does not then flow into domestic industries for the same reasons foreign capital does not move directly into these industries. One difference, however, is that foreign firms located within less developed countries have some advantage over outside investors in ascertaining profit opportunities in other fields. However, there is considerable immobility

of capital from such foreign-owned and foreign-directed firms into new product lines involving a very different technology from that used for existing production, especially if the optimum plant size is small. Foreign firms engaged, for example, in oil or copper production will vigorously seek out further profit opportunities in their own product lines, including those that establish additional forward and backward production linkages. But, a lack of interest and knowledge concerning the production and marketing of completely different products tends to offset their proximity advantage.

The flow of direct investment funds into developing countries involves not only an increase in the capital stock of these countries, but also an improvement in technology in the sectors affected. This means that, since foreign capital does not move into very many domestically-oriented industries, the state of technology in these industries remains backward. Thus, the explanation of why many less developed countries export capital-intensive products may rest on the immobility of capital between the export and domestic sectors and the technological disparity between these sectors compared to the same sectors in developed economies.

Although the analysis has dealt thus far with developed countries, it also has some applicability to resource-abundant countries like Canada. *U.S.* investment in natural resource industries in Canada tends to create a greater capital and technological disparity between export and domestically oriented industries than would exist without this investment. However, the experience of recent years has shown that as income and domestic markets grow in such developed countries, direct investment by the United States and other advanced countries takes place in product lines that formerly were mainly imported into these countries and are

characterized by significant scale economies. This seems to occur partly to take advantage of being located near the market and partly as a defensive response to import competing investments by domestic investors. To the extent that *U.S.* investment of this sort is in product lines that are more capital-intensive than other *U.S.* exports, the Leontief paradox tends to be reinforced.

One other matter that should be considered before presenting the empirical results of testing some of the different trade hypotheses is whether the Heckscher-Ohlin proposition should hold with respect to each pair of countries.¹⁶ Given a pure Heckscher-Ohlin-Samuelson model where all goods are traded, the number of products exceeds the number of factors, and factor-price equalization is achieved, the answer is that the proposition need not hold on a bilateral basis. When the number of commodities is greater than the number of productive factors, the precise distribution of world production and trade is indeterminate with a particular distribution of productive factors among countries and a given set of factor prices.¹⁷ It is not necessary, for example, for the most capital-abundant country to export a larger proportion of the total exports of the most capital-intensive product than a less capital-abundant country or indeed to export it at all. Within the limits set by factor prices, the actual pattern of intercountry production of any traded commodity depends upon a host of complex factors related to different historical rates of development. What is required in the Heckscher-Ohlin theory is simply that the capital-labor ratio of a capital-abundant country's total exports be greater than the

capital-labor ratio of its imports. It is quite possible for this relationship to hold with regard to a country's total trade but not with respect to its trade with a particular country.¹⁸ As Hodd (p. 22) has pointed out, preventing complete factor-price equalization by introducing transport costs into the two-factor model causes the Heckscher-Ohlin proposition to hold bilaterally as well as multilaterally, but this bilateral relationship can again break down when the model is complicated by additional factors, e.g., natural resources, that are complementary to one of the other factors, e.g., capital. Since Vanek and more recently Lawrence Weiser found evidence of a complementary relationship between capital and natural resources at least for the United States, the several empirical studies that have revealed inconsistencies in the factor-content of trade between two countries and the relative factor-endowment pattern of the two countries should not be regarded as providing evidence that necessarily runs counter to the Heckscher-Ohlin theory as it is now generally formulated.

II. Testing the Heckscher-Ohlin Theory and Other Trade Hypotheses

The major results of retesting the Heckscher-Ohlin hypothesis for the United States, using 1962 trade figures and 1958 capital, labor, and intermediate-input data are presented in Tables 1, 2, and 3. Table 1 presents factor-content (direct and indirect) ratios¹⁹ that compare representa-

¹⁸ See Appendix A for a geometric illustration of such a case. It is also not necessary under the Heckscher-Ohlin theory for the ratio of capital per worker in export industries to capital per worker in import-competing industries to be higher for a capital-abundant country than some less capital-abundant economy.

¹⁹ In testing the relationship between relative factor supplies and the factor content of trade, some writers (Keesing (1965) and Waehrer) compute only the direct factor content of exports and import replacements on the grounds that most intermediate inputs can be imported instead of produced domestically. This procedure confuses an *ex post* test of an equilibrium trade position

¹⁶ Bhagwati (pp. 175-76) raises this issue in his survey article and terms the lack of analysis on this point a serious deficiency in trade theory.

¹⁷ The net factor-trade balance is, of course, the same in these circumstances.

tive bundles of import competing with export products;^{20,21} Table 2 shows the distribution of the labor force by broad occupational groups; and Table 3 gives different regression estimates in which net adjusted trade balances²² of the various trading industries in the 1958 input-output table are made a function of different

to determine if the pattern of trade is consistent with the Heckscher-Ohlin theory with such exercises as predicting or planning for the detailed nature of a country's trade pattern, given its factor endowment and a set of international commodity prices. For the latter purpose the investigator must consider the possibility with respect to any possible export product that national income may be made greater by importing intermediate products rather than producing them domestically. Consequently, the optimum position may well be one where many intermediate inputs involved in trade are not produced locally. Lary's study of the potentialities for exports of manufactures in developing nations in which he analyzes only direct value-added ratios illustrates a problem where the use of direct factor-content ratios is the proper procedure. However, given a particular equilibrium pattern of trade, it is necessary to include both the direct and indirect labor and capital involved in producing exports and imports in order to determine a country's net trade balance in factor services via trade in commodities. If only direct coefficients are used, it is possible to conclude, for example, that a capital abundant country exports labor services and imports capital services when in fact it does the opposite. The direct coefficient test thus would erroneously infer that the Heckscher-Ohlin hypothesis failed to hold.

²⁰ The representative export and import-competing bundles do not include any services but instead are composed entirely of traded commodities.

²¹ Intermediate products imported and then reexported in the form of other products as well as imports containing intermediate inputs that were exported and then reimported should, of course, be excluded in calculating net factor flows. The Leontief method does in fact accomplish this since, for example, foreign produced intermediates that are imported and reexported are counted by the Leontief method on both the import and export side and thus net out in subtracting the factor services involved in exports from those involved in imports. If capital-labor ratios of exports and imports are compared, an incorrect ratio will be obtained but the error factor will not effect whether the quotient is above or below unity—which is the main purpose of the calculation.

²² Each industry's exports and competitive imports were adjusted by multiplying their respective share of total exports and competitive imports by one million dollars. An industry's net trade balance is the difference between these adjusted values for exports and competitive imports.

economic factors that allegedly influence the commodity pattern of trade.

One important result of the test is that the Leontief paradox still holds.²³ The ratio of capital per man-year in import-competing versus export production is 1.27²⁴ compared to the ratios of 1.30 and 1.06 that Leontief obtained for the 1947 and 1951 trade patterns, respectively.²⁵ Furthermore, in the various stepwise multiple regressions that were performed, the capital-labor ratio always entered first with a statistically significant negative sign as the single variable that best "explained" the trade pattern. However, if natural resource products are excluded, the

²³ Gary Hufbauer also obtains this result in his study of *U.S. trade in manufactures*.

²⁴ An estimate of this capital-labor ratio was also made in which transportation services, travel (weighted by an average of the capital-labor ratios for hotels and personal services, amusements, and miscellaneous manufactures) and other private services (weighted by an average of the capital-labor ratios for communications and radio and TV broadcasting) were included in the export and import competing bundles. The capital-labor ratio for imports rose to \$18,300 and that for exports to \$15,000. The import ratio divided by the export ratio was, therefore, 1.22.

²⁵ As Travis (pp. 98-99) has noted, the exclusion of noncompetitive imports from the *U.S.* import bundle because of the nonavailability in the United States of certain natural resources required for their production could conceivably result in an incorrect inference concerning the Heckscher-Ohlin hypothesis from a Leontief-type test. This would be the case if the production of noncompetitive imports was so highly labor-intensive (and would be so in the United States had the specific natural resources been available) that the capital-labor ratio of total imports, in contrast to just competitive imports, was less than the capital-labor ratio of exports. However, on the basis of a rough survey of the capital-labor ratios for noncompetitive imports produced abroad and given the fact that these imports constitute only about 8 percent of total *U.S.* commodity imports, it appears to be extremely unlikely that the labor intensity of noncompetitive imports could be so high as to account for the Leontief paradox. Actually, the capital-labor ratio calculated in the paper for competitive imports is so high that the capital-labor ratio for the 8 percent of imports (or 6 percent if traded services are included) which are noncompetitive would have to be negative in order to make the overall capital-labor ratio for imports even equal to the capital-labor ratio for exports.

TABLE 1—FACTOR REQUIREMENTS (DIRECT AND INDIRECT) PER MILLION DOLLARS OF U.S. EXPORTS AND COMPETITIVE-IMPORT REPLACEMENTS, 1962

	Imports	Exports	Import/ Export Ratio
Net Capital			
All Industries	\$2,132,000	\$1,876,000	1.14
Excl. Agriculture	1,806,000	1,403,000	1.29
Excl. Natural Resource Products ^a	1,259,000	1,223,000	1.03
Gross Capital			
All Industries	\$2,393,000	\$2,196,000	1.09
Excl. Agriculture	2,083,000	1,777,000	1.17
Excl. N. R.	1,582,000	1,599,000	.99
Labor (man-years)			
All Industries	119	131	.91
Excl. Agriculture	100	109	.92
Excl. N. R.	106	107	.99
Net Capital-Labor			
All Industries	\$18,000	\$14,200	1.27
Excl. Agriculture	18,100	12,800	1.41
Excl. N. R.	11,900	11,500	1.04
Average Years of Education of Labor			
All Industries	9.9	10.1	.98
Excl. Agriculture	10.2	10.6	.96
Excl. N. R.	10.3	10.7	.97
Average Costs of Education of Labor			
All Industries	\$10,300	\$10,500	.97
Excl. Agriculture	11,000	11,900	.92
Excl. N. R.	11,200	12,200	.92
Net Capital Plus Total Cost of Education ÷ Labor			
All Industries	\$28,300	\$24,700	1.14
Excl. Agriculture	29,100	24,700	1.18
Excl. N. R.	23,100	23,700	.97
Average Earnings of Labor			
All Industries	\$4,570	\$4,660	.98
Excl. Agriculture	5,050	5,460	.92
Excl. N. R.	5,030	5,400	.93
Proportion of Engineers and Scientists			
All Industries	.0189	.0255	.74
Excl. Agriculture	.0230	.0352	.65
Excl. N. R.	.0228	.0369	.62
Scale Index			
All Industries	51	56	.91
Excl. Agriculture	55	66	.83
Excl. N. R.	57	67	.85
Unionization Index			
All Industries	59	62	.95
Excl. Agriculture	65	72	.90
Excl. N. R.	65	71	.92
Concentration Index			
All Industries	39	40	.98
Excl. Agriculture	42	46	.91
Excl. N. R.	41	46	.89
Proportion of Labor with			
0-8 years of education	.39	.37	1.05
9-12 years of education	.49	.50	.98
13+ years of education	.12	.13	.92

(Footnote ^a will be found on next page.)

capital-labor ratio falls to 1.04 when capital is measured on a net basis and to 1.00 when capital is measured in gross terms.

The hypothesis that export production involves higher skill requirements than import competing production also receives support, as the figures on average earnings, average years of education, and average costs of education indicate. A crude measure of the amount of physical and human capital used in export versus import-competing production was calculated by combining the data on physical capital and the costs of education.²⁶ As Table 1

²⁶ To obtain direct education costs, the figures on years of education from the 1960 sample census, supplemented with data on school retention rates, were

indicates, adding this measure of human capital to the physical capital figure is not sufficient to reverse the Leontief results for all industries combined but does reverse it when natural resource industries are excluded.

Classifying the labor force involved in export and import competing production by levels of education and by various occupational groups further brings out the importance of the skill factor in explaining U.S. trade. The educational breakdown

multiplied by the 1956 cost figures determined by Theodore Schultz (p. 34). Estimates of foregone earnings were added to these direct costs to obtain total education costs. However, no measure of accumulated interest costs is included in the estimate nor does it include any on-the-job training costs.

*Natural resource products were arbitrarily defined as all agricultural and mining industries (1-10); tobacco manufactures (15); lumber and wood products (20); petroleum refining (31); and primary nonferrous metals manufacturing (38). The list is roughly similar to the one used by Leontief except that petroleum refining is added and non-livestock agricultural products are included. One could argue quite persuasively that other industries should also be included.

Sources: The coefficients of total requirements (direct and indirect) per dollar of delivery to final demand were taken from U.S. Department of Commerce. The employment figures used to calculate the 1958 direct labor coefficients for the 79 industries covered in the study were furnished by Jack Alterman of the Bureau of Labor Statistics.

The 1958 capital coefficients for industries 12-64 were obtained by reclassifying data given in *Census of Manufactures, 1958*. Net capital is the sum of net book value, work in progress, materials, and finished product inventories. Finished product inventories were adjusted to purchasers prices and, by utilizing the transaction matrix of the input-output table, were distributed to the industries using them. The gross capital coefficients are based on gross book value rather than net book value. The coefficients for the non-manufacturing sectors are based on a wide variety of sources. They include the two basic Leontief articles (1953 and 1956); the study by Leontief et al.; Bert G. Hickman; John W. Kendrick; and Daniel B. Creamer, Sergei P. Dobrovolsky and Israel Borenstein.

Export and import data for the 60 commodity sectors in which trade occurred in 1962 are from *Exports and Imports as Related to Output*. Values of exports and imports for 1962 were adjusted to 1958 prices by deflators that in the case of the manufacturing sectors (13-64) were provided by the Office of Business Economics,

Department of Commerce and that for mining and agriculture were obtained from the 1962 *Minerals Yearbook* and *Wholesale Prices and Price Indexes*, 1962. Imports were multiplied by ratios of landed value to foreign port value and exports by ratios of producer value to export value in order to make them comparable to the producer-value figures of the input-output table. These ratios were provided by the Office of Business Economics, Department of Commerce.

The Bureau of the Census, *1/1000 Sample of the Population of the United States, 1960* was available in data tape form from the Social Systems Research Institute computation library at the University of Wisconsin. Using the industrial classification system employed by the Census, it is possible to arrange the labor force covered in the census into the same industry groups adopted for the 1958 input-output table. Characteristics relating to years of education, occupations, and earnings were then determined from the data tape for these individuals.

The direct scale and unionization ratios were adapted from Leonard Weiss. The direct scale index for each industry is based on the percentage of employees in establishments with 250 or more employees, and the direct unionization index represents the percentage of an industry's production workers employed in plants where a majority of the workers are covered by collective bargaining contracts. The direct concentration figures, also from the paper by Weiss, are four-firm concentration ratios, adjusted for the local or regional character of certain industries. The row vector of direct ratios for each of the three variables was postmultiplied by the inverse matrix, and weighted averages of the direct and indirect requirements for these characteristics then obtained for each industry by dividing the resulting row vector by the appropriate column sums of the inverse matrix.

TABLE 2—DISTRIBUTION OF LABOR FORCE BY SKILL GROUPS, PER MILLION DOLLARS OF EXPORTS AND COMPETITIVE-IMPORT REPLACEMENTS, 1962

(in percentages)

(A) Six Skill Groups	Im-ports	Ex-ports	Import/ Export Ratio
I. Professional, technical and managerial	12.0	12.5	.96
II. Clerical and sales	15.2	15.1	1.01
III. Craftsmen and foremen	14.9	15.4	.97
IV. Operatives	30.4	25.1	1.21
V. Laborers (nonfarm) and service	10.3	7.5	1.37
VI. Farmers and farm laborers	17.2	24.4	.70
	100.0	100.0	
(B) Eleven Skill Groups			
I. Professional and technical	5.7	6.7	.85
II. Managerial, except farm	6.3	5.8	1.09
III. Craftsmen and foremen	14.9	15.4	.97
IV. Sales	4.4	4.1	1.07
V. Clerical	10.8	11.0	.98
VI. Operatives	30.4	25.1	1.21
VII. Laborers, except farm	6.9	4.3	1.60
VIII. Service, except private household	3.1	2.9	1.07
IX. Farmers and farm managers	11.2	15.8	.71
X. Private household workers	.3	.3	1.00
XI. Farm laborers and foremen	6.0	8.6	.70
	100.0	100.0	

Sources: See Table 1.

indicates that the proportions of individuals with 9–12 years of education and especially with 13 or more years of education are higher in export than in import competing production, whereas the share of those with only 0–8 years of education is higher on the import side. As the occupational figures (Table 2) show, farmers and farm laborers, who are among the least educated occupational groups, are considerably more important in export

than import competing production. However, nonfarm laborers and operatives, who are also at the lower end of the educational attainment scale, are sufficiently more important in import competing production compared to export activities to make the proportion of the labor force as a whole with only a primary school education more significant in import competing than export production. The other occupations that stand out as more significant on the export than import competing side are professional, technical and managerial employees and craftsmen and foremen. Clerical and sales employees do not differ in their relative importance in export versus import competing activities.

The correlation analysis (Table 3, Equations 1a, 2, and 3) shows that there is a significant positive relationship between the percentage of engineers and scientists, craftsmen, and farmers in an industry and the net world export surplus of the industry. The percentage of operatives and nonfarm laborers have the expected negative signs but the coefficients are not significant. As regression equation 7 indicates, there is also a statistically significant positive relationship between the importance in an industry of those with more than a high school education and the industry's world trade balance.

Research and development activities also show up as being much more important in export output than in import-competing production. The ratio of the R&D costs involved in producing a representative bundle of import-competing versus export commodities, as calculated from the R&D sector in the input-output table, is .66.²⁷ The ratio of the number of engineers and natural scientists engaged in import-competing versus export activities

²⁷ The R&D sector in the input-output table includes, however, only research and development performed for sale and thus excludes R&D performed within a company.

TABLE 3—REGRESSION EQUATIONS RELATING SELECTED ECONOMIC CHARACTERISTICS BY INDUSTRY TO U.S. WORLD AND BILATERAL TRADE BALANCES BY INDUSTRY

Independent Variables													
Dependent Variable (in dollars)	Capital-Labor	Cap. Plus Costs of Educ.- Labor	Percentage of the Labor Force in Various Skill Groups						Scale Index	Union- ization Index	Conc. Index	R ²	
			Eng. & Sc.	Rest of I	II	III	IV	V					VI
1a. Adjusted Exports Minus Adj. Imports ^a (World)	-1.37 -4.35**	—	7011 2.13*	-1473 .69	71 .06	1578 1.96*	-248 .79	-761 .89	845 3.81**	-421 -1.25	343 1.11	—	.44 3.85**
1b. <i>X-M</i> (Canada)	-1.38 -2.39*	—	-4608 .76	1778 .45	512 .22	1127 .76	-335 .58	-2771 -1.81*	546 1.35	302 .49	2 .00	—	.24 1.52
1c. <i>X-M</i> (W. Europe)	-.27 -.68	—	8185 1.98*	1397 .52	-1040 .66	907 .90	34 .09	-958 .91	1095 3.93**	-983 -2.32*	389 1.00	—	.32 2.31*
1d. <i>X-M</i> (Japan)	-.06 -.13	—	8748 1.74*	2242 .68	-930 .48	1010 .82	-718 -1.49	-2429 -1.90*	1333 3.93**	-1381 -2.68**	1047 2.22*	—	.42 3.51**
1e. <i>X-M</i> (LDCs)	-2.73 -5.80**	—	-768 .16	-1373 .43	-948 .51	2701 2.25*	-315 .67	435 .35	659 1.99*	255 .51	220 .48	—	.51 4.99**
1f. <i>X-M</i> (Others)	-1.11 -1.49	—	27496 3.54**	-13335 -2.64*	3761 1.27	2582 1.36	-15 .02	3041 1.54	-402 -.77	-201 -.20	-506 -.70	—	.37 2.84**
2. <i>X-M</i> ^a (World)	-1.20 -3.98**	—	5789 1.97*	-1145 .52	-84 .07	1603 1.97*	-310 -1.01	-728 .86	854 3.75**	—	295 .97	-478 -1.03	.44 3.76**
3. <i>X-M</i> ^a (World)	— -4.38**	-1.36 -4.38**	7631 2.30*	-1156 .54	298 .24	1663 2.05*	-184 .58	-620 .74	928 4.05**	-406 -1.21	345 1.12	—	.45 3.88**
Dependent Variable	Constant	Capital-Labor	Cap. Plus Costs of Educ.- Labor	Percentage of Labor Force With Various Yrs. of Educ.			Percent Sc. & Eng.	Av. Costs of Educ.	Av. Yrs. of Educ.	Scale Index	Union- ization Index	Conc. Index	R ²
				1-8 yrs.	9-12 yrs.	13+ yrs.							
4. <i>X-M</i> (World)	65256 1.66	-1.09 -3.29**	—	—	—	—	10080 2.25*	-6.64 -1.50	—	-319 -.86	259 .90	—	.25 3.45
5. <i>X-M</i> (World)	108312 1.16	-1.12 -3.32**	—	—	—	—	—	—	-10728 -1.06	-305 -.81	249 .85	—	.23 3.16**
6. <i>X-M</i> (World)	17246 1.02	— -3.26**	-1.04 -3.26**	—	—	—	4274 1.67	—	—	—	81 .31	-157 -.31	.21 3.60**
7. <i>X-M</i> (World)	—	-1.35 -3.86**	—	640 -1.57	-844 -1.46	3090 2.00*	—	—	—	—	—	—	.22 3.83**

Sources: See Table 1

^a The constant term has been suppressed in equations 1, 2, 3, and 7 because of the indeterminacy resulting from the fact that the skill-level and education-level percentages add to unity in each industry.* and ** indicate 90 percent and 99 percent significant levels for the *t* values (shown in italics) of the regression coefficients and for the *F*-ratio of the squared multiple correlation coefficient.

is .74.²⁸ Moreover, as already noted, in the regression model used, the percentage or the absolute number of engineers and scientists in an industry appears as a significant variable that is positively correlated with the industry's export surplus. This relationship is especially strong when natural resource products are excluded from the trade pattern.

Another exercise confirming the importance of this variable is the correlation between the percentage change of exports in each industry from 1947 to 1962 and various characteristics of the labor force in each industry, such as their earnings, years of education, a simple skill index, the absolute importance of engineers and scientists as well as general industry characteristics such as the degree of concentration, unionization, and large scale employment.²⁹ The engineers-scientists variable and the concentration ratio are the only two variables that are significantly related (positively) to the growth of exports. When the same variables are used to explain the percentage change in imports, none comes out to be statistically significant.

²⁸ This group includes both individuals engaged in research and development as well as those engaged in current production activities. Using data for eighteen industries and direct requirements only, Kenen (1968) compared the relative importance of the two groups in "explaining" trade patterns and obtained ambiguous results. As Keesing (1968 pp. 175-89) had previously shown, for exports alone, the ratio of scientists and engineers engaged in research and development to the total labor force in the industry is statistically significant whereas the proportion of scientists and engineers in non-R&D activities is not. On the other hand, when an industry's net trade balance is taken as the dependent variable, the opposite result is obtained. Consequently, in view of these results and also because of the similarity between the results obtained in this study from using R&D expenditures and the number of scientists and engineers, it seems best to regard the engineers-scientists variable used here as both a skill measure and a proxy for R&D activities that result in new and improved products.

²⁹ In this exercise the 1947 trade data were classified on the basis of the industrial breakdown in the 1958 input-output table. The various economic characteristics of the industries pertain to the period around 1960.

Beside determining the trade requirements for engineers and scientists, an estimate was made of the import versus export requirements of top management, i.e., managers, officials, and proprietors earning more than \$10,000. Their numbers are larger in export production than in import-competing production but the proportions of the total labor force engaged in top management activities are about the same on the import and export sides.³⁰

Two other industry characteristics of special interest are the relative importance of scale economies and the degree of unionization in import-competing versus export production. Steffan Linder has stressed the point that profitable production for home markets is a necessary condition for manufactured products to be potential export products. Consequently, in industries where scale economies are important, the size of the American market may give the United States an export advantage, quite aside from any other factors. The unionization calculation is aimed at the hypothesis that unions may raise wages above their competitive levels and thus act to offset underlying "real" factors that contribute to a country's comparative advantage.³¹ A variable reflecting the degree of industrial concentration is also introduced into the analysis, but it is highly correlated ($r = .87$) with the scale index. As Table 1 indicates, the scale factor, the degree of unionization, and the degree of concentration are all more important for export production than for import-competing activities. However, none of these variables turn out to be sta-

³⁰ Two other results that may be of interest are that there is no difference between import-competing and export production in the average age of the workers or the proportion who are white.

³¹ Even assuming that unions do raise wages above competitive levels, this hypothesis depends on the assumption that the resulting competitive disadvantage is not offset by the same force operating in other countries.

tistically significant in the regression analyses of total trade.

In order to indicate the effect of import duties as well as some of the main non-tariff trade barriers on the capital-labor ratio employed in import-competing production, import demand elasticities were assigned to the various trading industries in the input-output table and a new per million dollar bundle of competitive-imports was then determined under the assumption that the average duty (or the ad valorem duty-equivalent of the non-tariff barrier) in each industry was reduced to zero.³² The fact that the capital-labor ratio with the new commodity-composition of imports is about 5 percent lower than the ratio computed with the actual import bundle confirms Travis's contention that tariffs operate in the direction of the Leontief paradox.³³ Furthermore, the commercial policies of other countries probably tend to reduce the average capital-labor ratio in export production below its free trade level. A removal of all trade-distorting measures might confirm the expectation of the Heckscher-Ohlin model for the United States, but, based on my own study of tariff and nontariff barriers to trade, I do not think that this would be the case. However, further empirical study of this subject is very much needed.

The effect of an increase in income levels on the capital-labor ratio of import com-

peting production was also estimated by assigning appropriate income elasticities of import demand to various commodity groups.³⁴ The capital-labor ratio required to produce the increment in imports associated with an increase in income is \$17,750 or slightly less than the \$18,000 average for the 1962 bundle of competitive imports. Thus, if the *U.S.* commodity structure of income elasticities of import demand is typical of the pattern for the rest of the world towards the United States, then demand differences related to income differences among countries are not a factor that tends to account for the Leontief paradox.

In addition to determining the factor content of a representative bundle of *U.S.* exports to all countries as a whole and a representative bundle of competitive imports from the rest of the world, the Heckscher-Ohlin hypothesis was tested with respect to *U.S.* trade vis-à-vis Western Europe, Japan, Canada, and less developed countries, and an all other group (mainly Oceania).³⁵ As previously noted, the assumptions necessary for the Heckscher-Ohlin proposition to be logically true with regard to a country's total

³² The indirect effect that reducing duties on products used as intermediate inputs has in increasing domestic production and thus reducing imports was not taken into account in estimating the new import bundle.

³⁴ The income elasticities as well as the commodity classification employed in making the estimate are from the Ball-Marwah article. The particular income elasticities used were: crude foodstuffs .49; manufactured foodstuffs .96; crude materials .87; semi-manufactures 1.22; and manufactured goods 2.47.

³⁵ Western Europe consists of all European members of OECD; the less developed countries are composed of other Asia (Asia except for Japan and China Mainland), Africa, and other America (Americas excluding the United States and Canada); and the all other group is made up of Eastern Europe, other Europe, China Mainland, and Oceania. The data on which these regional trade patterns are based are much less detailed than those from which the world trade pattern is derived. Furthermore, the fact that no products are excluded from the export side on the basis of being non-competitive is a more serious drawback for bilateral analyses than for the analysis of *U.S.* trade with the rest of the world.

³³ Using the study by R. J. Ball and K. Marwah, industries listed in the input-output table were divided into five groups and assigned the following import-demand elasticities: crude foodstuffs -.46; manufactured food-stuffs -2.39; crude materials -.38; semi-manufactured products -1.64 and manufactured goods -4.04. Tariff rates for 1962 were obtained by dividing the calculated import duty for an industry by the value of its imports. The nontariff barriers included were the quotas on agricultural products, cotton textiles, and petroleum as well as the American Selling Price system of valuing certain chemicals. See Baldwin (p. 163). In estimating the price effect of the cuts in the degree of protection, it was assumed that the elasticity of foreign import-supply was infinite for all commodities.

TABLE 4—FACTOR-CONTENT RATIOS FOR U.S. BILATERAL TRADE PATTERNS WITH SELECTED REGIONS, 1962

Economic Characteristic	Import/Export Ratios ^a									
	Canada		Western Europe		Japan		LDCs		Others	
	All Sectors	Excl. N.R.	All Sectors	Excl. N.R.	All Sectors	Excl. N.R.	All Sectors	Excl. N.R.	All Sectors	Excl. N.R.
Net Capital	1.25	1.02	.83	1.01	.77	1.03	1.44	1.05	1.52	1.09
Gross Capital	1.16	.99	.87	.99	.83	1.03	1.28	.93	1.40	1.04
Labor	.89	.88	.95	1.07	1.05	1.23	.81	.92	1.07	.82
Net Capital-Labor	1.41	1.15	.87	.93	.73	.84	1.78	1.14	1.42	1.33
Average Years of Education	.99	1.01	1.02	.98	1.01	.95	.98	.96	.92	.95
Average Costs of Education	.97	.99	1.05	.94	1.03	.88	.96	.89	.81	.87
Net Capital Plus Cost of Education÷Labor	1.19	1.07	1.19	.89	1.22	.87	1.24	1.01	1.64	1.06
Average Earnings	.97	.91	1.05	.94	1.04	.88	.92	.81	.80	.90
Proportion of Engineers and Scientists	.82	.90	1.04	.74	.94	.64	.75	.47	.29	.37
Scale Index	.77	.86	1.27	1.07	1.28	1.07	.70	.68	.58	.62
Unionization Index	.86	.95	1.20	1.05	1.14	.99	.77	.81	.77	.83
Concentration Index	.83	.85	1.20	1.05	1.10	.96	.82	.78	.75	.77
Skill Group I ^b	.96	.94	1.10	.87	1.08	.77	.93	.83	.66	.81
II	.89	.96	1.18	.96	1.16	.88	.95	1.04	.75	1.13
III	.91	.89	1.20	.92	1.12	.81	.83	.71	.62	.89
IV	1.04	1.07	1.42	1.14	1.67	1.32	1.03	1.23	.82	1.12
V	1.60	1.13	1.21	1.02	1.44	.95	1.12	1.06	1.24	1.44
VI	.82	—	.46	—	.24	—	1.11	—	.24	—

Sources: Bilateral trade data are from OECD, *Foreign Trade Statistical Bulletins*. See Table 1 for sources of other data.

^a Import/Export ratios are computed in the same manner as in Table 1.

^b The skill groups are the same as in Table 2.

trade do not imply that the theory must hold on a bilateral basis. However, a regional analysis is useful in revealing additional information on the factors influencing the commodity pattern of U.S. trade. The results of these tests (Table 3, equations 1b-1f and Table 4) are that the Leontief result does not hold with respect to either U.S.-Western European or U.S.-Japanese trade but does exist with respect to trade between the United States and Canada, the United States and less developed countries, and the United States and all other countries. The latter three groups of countries represent regions that are relatively abundant in natural resources. In view of the strong complementary relationship between capital and certain natural resources and the previously made point concerning the inter-

national flow of U.S. capital, technology, and top management into export-oriented, natural resource industries in foreign countries, the results with regard to Canada, the less developed countries, and Oceania are not unexpected.³⁶ The trade

³⁶ As Table 4 indicates, imports from these regions are more capital-intensive than exports to them even when the list of natural resource industries are excluded from the calculations and when human capital is added to physical capital. However, the nature of much of the remaining trade is still greatly influenced by transportation and technical processing considerations that favor location of production in these areas. Such is the case, for example, in the very important food sector (14) where imports from the LDCs are dominated by cane sugar and imports from the all other group by meat as well as for the paper industry (24) where imports of pulp and newsprint from Canada are large. When these two industries are also excluded from the factor content calculations, the ratio of physical plus human capital to labor is lower in import competing than export production for all regions, and the ratio of physical capital

patterns with Western Europe and Japan are not as heavily influenced by imports of natural resource products nor is direct foreign investment as important in these regions relative to domestic capital accumulation. Thus, relative domestic supplies of capital and labor play a more important role in determining the trade structure between the United States and these regions.

Although the United States exports comparatively capital-intensive products to Western Europe and Japan, the capital-labor variable does not show up as statistically significant in the multiple regression analysis with respect to these areas. Other factors appear to be more important as determinants of the trade patterns between the United States and these regions. In particular, there is a significant positive relationship between the percentage of engineers and scientists employed in an industry and the industry's net export balance with respect to each of these two regions (as well as for the all other group). The scale variable also shows up as significant for these two regions. Rather surprisingly, the sign of the scale coefficient is negative (largely because of the large export surplus of agricultural products to Western Europe and Japan).³⁷

General measures of skill and human capital such as average costs of education, average years of education, and average earnings were found to be statistically significant only in the case of *U.S.-Japanese* trade.³⁸ However, the percentage of employees with 13 or more years of education required to produce a given value of output in each industry is significantly

correlated (positively) with an industry's net trade balance between the United States and Western Europe, the United States and Japan, and the United States and the all other group. A similar positive correlation holds for *U.S.-Western European* trade with regard to the proportion of the labor force educated 8 years or less, whereas a significant negative relationship exists for *U.S.-Western European* and *U.S.-Japanese* trade with regard to those receiving 9–12 years of education. Dividing the labor force into broad occupational groups further reveals the importance of various types of labor skills in a manner generally consistent with what one would expect from a factor-proportions approach. The significance of the engineers-scientists variable has already been mentioned. The number³⁹ and percentage of unskilled (nonfarm) workers employed in an industry are significantly correlated (negatively) with the industry's export surplus in the cases of Japan and Canada, and the number (though not percentages) of semi-skilled workers enters significantly with the same negative sign for Japan. Furthermore, the number and percentage of skilled workers in an industry is significantly correlated in a positive manner with the industry's net balance of trade between the United States and the less developed countries.

III. Conclusions

The preceding analysis strongly supports the view that a straightforward application of a two-factor (capital and labor) factor-proportions model along Heckscher-Ohlin lines is inadequate for understanding the pattern of *U.S.* trade. Not only is the sign of the capital-labor ratio opposite from what would be ex-

to labor is lower in import competing than export production for all regions except the less developed countries.

³⁷ When natural resource products are eliminated, the scale factor is not statistically significant for Western Europe or Japan.

³⁸ These results as well as those that are discussed in the next two sentences are not reported in Table 3.

³⁹ A regression equation (not shown in Table 3) was estimated in which the independent variables were the capital-output ratio for each industry and the number of employees in each of the six skill groups.

pected from the model but it is statistically significant in this unexpected direction. What this negative sign seems to reflect is, as Vanek and others have suggested, that there is a strong complementarity between certain natural resources—many forms of which are relatively scarce in the United States—and physical capital.⁴⁰ The regional breakdown indicates, for example, that the source of the paradox is the pattern of *U.S.* trade with Canada, the less developed countries, and the all other group—all of which export a significant volume of natural resource products to the United States. When various natural resource products are eliminated from the factor-content calculations, the overall ratio of capital per worker in import-competing production to capital per worker in export production drops from 1.27 to 1.04. Omitting natural resource industries from the regression data also eliminates the capital-labor ratio as a statistically significant variable. Moreover, in the remaining group there are still some important industries in which the costs of transporting natural resource products required as inputs are relatively high and whose location, therefore, tends to be near the source of the natural resources that are indirectly required for production.

As previously noted, the complementary relationship between physical capital and natural resources need not be offset by *U.S.* exports of goods that are even more capital-intensive than natural resource products or by other imports of a highly labor-intensive nature, provided the various assumptions of a simple Heckscher-Ohlin model with respect to capital mobility, homogeneity of the labor supply, commercial policy, and technological parity do not hold. Evidence indicating that

the capital mobility assumptions do not hold is available in the economic development literature whereas data consistent with the position that the labor-supply, technology, and commercial-policy assumptions of the traditional model do not hold have been presented in this paper. It seems clear from the preceding analysis that the relatively abundant supply of engineers and scientists is an important source of the United States' comparative-advantage position, especially as far as trade in manufactures is concerned. This abundance of highly trained labor gives the United States an export advantage in products requiring relatively large amounts of such labor.⁴¹ Probably of even more importance is the fact that a significant part of this labor group is engaged in research and development activities. Even those working directly in production facilitate the development of product improvements. Thus, in product lines where the technological opportunities for product improvements are favorable, the use of engineers and scientists for research and development activities fosters temporary *U.S.* trade advantages based on technological differences rather than on relative factor proportions. Just how to weigh the relative importance of these two aspects of the engineers-scientists variable cannot be determined from this study but what evidence there is suggests that both are significant in influencing the pattern of *U.S.* trade.⁴²

The relative supplies of certain other types of labor skills also appear to be important determinants of the structure of *U.S.* commodity trade. As would be expected in a Heckscher-Ohlin model with

⁴⁰ As Seiji Naya (p. 567) has pointed out, this complementary relationship has a more general applicability among countries when agricultural products are excluded from the list of natural resource products.

⁴¹ In most industries engineers and scientists make up only a small fraction of the labor force. However, in 1960 the proportion was between 5 and 10 percent for thirteen of the sixty input-output industries in which international trade took place.

⁴² Kenen (1968) concludes from his analysis of this matter that an eclectic approach is still in order.

several types of labor, the United States not only indirectly exports professional and technical labor but also skilled craftsmen and foremen. Furthermore, we indirectly import semi-skilled and unskilled (nonfarm) labor, both of which are usually considered to be comparatively scarce in the United States.

General measures of human capital such as earnings, years of education, and costs of education, fail to capture much of the explanatory power that is given by a breakdown into levels of educational attainment or into traditional skill groups. Part of this may be due to the omission of on-the-job training from the general education variable used in the study. The positive correlation between net exports and average years of education is also weakened by the large export surplus in the agricultural sector, where formal educational requirements are not only low but where the relatively abundant supply of land in the United States plays a significant export-creating role. The existence of social and economic institutional arrangements that impede equilibrating educational adjustments within the labor force and rapid technological progress that frequently changes educational requirements further tend to diminish the statistical significance of gross measures of the stock of human capital as explanatory variables of the *U.S.* trade pattern. Simple (and imperfect) measures of the degree of scale economies, unionization, and industrial concentration also do not account for the *U.S.* trade pattern in a statistically significant manner. Protection levels in each industry, on the other hand, do seem to influence the pattern and factor-content of trade to an appreciable extent.

The clearest conclusion to be drawn from the study is, as other writers (see articles by Kenen (1968) and Hufbauer) have emphasized recently, that it is necessary to discard simple, single-factor (e.g.

capital per worker) trade theories in favor of multi-factor trade models. In particular, the labor force must be divided into various skill groups and the notion of relative differences in human capital taken into account. Other variables, such as natural resource conditions, technological differences, transportation costs, and commercial policies, must be explicitly included in these models. Furthermore, trade theory should take greater account of the degree of difficulty with which productive factors move among sectors within an economy and especially barriers to the flow of factors abroad into various sectors of an economy. Under this more general approach the relative abundance among countries of the factors of production will still occupy an important place in trade theory but a more complex notion of productive factors will be utilized and other considerations will also play important explanatory roles. Moreover, as we improve our predictive powers with these broader trade models, we must devote greater efforts to understanding the processes that determine the nature of the underlying variables affecting trade patterns. This, in turn, should enable us to construct a more fundamental, dynamic theory of international trade.

APPENDIX A

The point that the Heckscher-Ohlin theorem does not necessarily imply that the Heckscher-Ohlin relationship must hold bilaterally, given a standard model with two or more factors, a greater number of products than factors, and at least three countries, can be illustrated with Figure 1. The points *X*, *Y*, and *Z* represent the labor and capital endowment of countries *X*, *Y*, and *Z*, respectively. For simplifying purposes it is assumed that the equilibrium production pattern is such that Country *X* produces only commodity *A* (*OX* of *A*), Country *Z* produces only commodity *C* (*OZ* of *C*), whereas Country *Y* produces *Ok* of *C* plus *km* of *A*,

- Economic Theory: Growth and Development*, New York 1965, 2, 173-5.
- R. Bharadwaj, "Factor Proportions and the Structure of Indo-U.S. Trade," *Indian Econ. J.*, Oct. 1962, 10, 105-16.
- and J. Bhagwati, "Human Capital and the Pattern of Foreign Trade: The Indian Case," *Indian Econ. Rev.*, Oct. 1967, 2, 117-42.
- G. Bickel, *Factor Proportions and Relative Price Under CES Production Functions: An Empirical Study of Japanese-U.S. Comparative Advantage*, Stanford 1966.
- A. J. Brown, "Professor Leontief and the Pattern of World Trade," *Yorkshire Bull. Econ. Soc. Res.*, Nov. 1957, 9, 63-75.
- D. B. Creamer, S. P. Dobrovolsky, and I. Borenstein, *Capital in Manufacturing and Mining*, Princeton 1960.
- P. David and Th. van de Klundert, "Biased Efficiency Growth and Capital-Labor Substitution in the U.S., 1899-1960," *Amer. Econ. Rev.*, June 1965, 55, 357-94.
- M. A. Diab, *The United States Capital Position and the Structure of Its Foreign Trade*, Amsterdam 1956.
- P. T. Ellsworth, "The Structure of American Foreign Trade: A New View Examined," *Rev. Econ. Statist.*, Aug. 1954, 36, 274-85.
- J. L. Ford, *The Ohlin-Heckscher Theory of the Basis and Effects of Commodity Trade*, New York 1965.
- W. H. Gruber and R. Vernon, "The R & D Factor in a World Trade Matrix," Univ.-Nat. Bur. Comm. Econ. Res., Conference on Technology and Competition in World Trade, New York Oct. 11-12, 1968.
- G. Hanoch, "An Economic Analysis of Earning and Schooling," *J. Hum. Resources*, summer 1967, 2, 310-29.
- K. E. Hansson, "A General Theory of the System of Multilateral Trade," *Amer. Econ. Rev.*, Mar. 1952, 42, 59-68.
- B. G. Hickman, *Investment Demand and U.S. Economic Growth*, Washington 1965.
- M. Hodd, "An Empirical Investigation of the Heckscher-Ohlin Theory," *Economica*, Feb. 1967, 34, 20-29.
- G. C. Hufbauer, "The Commodity Composition of Trade in Manufactured Goods," Univ.-Nat. Bur. Comm. Econ. Res., Conference on Technology and Competition in International Trade, New York, Oct. 11-12, 1968.
- D. Keesing, "The Impact of Research and Development on United States Trade," in P. Kenen and R. Lawrence, eds., *The Open Economy: Essays on International Trade and Finance*, New York 1968, 175-89.
- , "Labor Skills and Comparative Advantage," *Amer. Econ. Rev. Proc.*, May 1966, 56, 249-55.
- , "Labor Skills and International Trade: Evaluating Many Trade Flows with a Single Measuring Device," *Rev. Econ. Statist.*, Aug. 1965, 47, 287-94.
- , "Labor Skills and the Structure of Trade in Manufactures," in P. Kenen and R. Lawrence, eds., *The Open Economy: Essays on International Trade and Finance*, New York 1968, 3-18.
- J. W. Kendrick, *Productivity Trends in the United States*, Princeton 1961.
- P. Kenen, "Nature, Capital and Trade," *J. Polit. Econ.*, Oct. 1965, 73, 437-60.
- , "Skills, Human Capital and Comparative Advantage," Univ.-Nat. Bur. Comm. Econ. Res., Conference on Human Resources, Madison, Wisc., Nov. 16, 1968.
- I. Kravis, "Wages and Foreign Trade," *Rev. Econ. Statist.*, Feb. 1956, 38, 14-30.
- M. Kreinin, "Comparative Labor Effectiveness and the Leontief Scarce Factor Paradox," *Amer. Econ. Rev.*, Mar. 1965, 55, 131-40.
- H. B. Lary, *Imports of Labor-Intensive Manufactures from Less Developed Countries*, New York 1968.
- W. Leontief, "Domestic Production and Foreign Trade: The American Capital Position Re-examined," *Proc. of the Amer. Philosophical Soc.*, Sept. 1953, 97, 332-49.
- , "Factor Proportions and the Structure of American Trade: Further Theoretical and Empirical Analysis," *Rev. Econ. Statist.*, Nov. 1956, 38, 386-407.
- , "International Factor Costs and Factor Use," *Amer. Econ. Rev.*, June 1964, 54, 335-45.
- et al, *Studies in the Structure of the American Economy*, New York 1963.

- S. Linder, *An Essay on Trade and Transformation*, New York 1961.
- B. S. Minhas, *An International Comparison of Factor Costs and Factor Use*, Amsterdam 1963.
- S. Naya, "Natural Resources, Factor Mix, and Factor Reversal in International Trade," *Amer. Econ. Rev. Proc.*, May 1967, 57, 561-70.
- B. Ohlin, *Interregional and International Trade*, Cambridge 1933.
- G. Philpot, "Labor Quality, Returns to Scale and the Elasticity of Factor Substitution," *Rev. Econ. Statist.*, May 1970, 52, 194-99.
- K. W. Roskamp, "Factor Proportions and Foreign Trade: The Case of West Germany," *Weltwirtschaftliches Archiv*, 1963, 91, 319-26.
- K. Roskamp and G. McMeekin, "Factor Proportions, Human Capital and Foreign Trade: The Case of West Germany Reconsidered," *Quart. J. Econ.*, Feb. 1968, 82, 152-160.
- T. Schultz, *The Economic Value of Education*, New York 1963.
- M. Tatemoto and S. Ichimura, "Factor Proportions and Foreign Trade: The Case of Japan," *Rev. Econ. Statist.*, Nov. 1959, 41, 442-6.
- W. P. Travis, *The Theory of Trade and Protection*, Cambridge 1964.
- J. Vanek, "The Factor Proportions Theory: The N-Factor Case," *Kyklos*, Oct. 1968, 21, 749-56.
- , *The Natural Resource Content of United States Foreign Trade, 1870-1955*, Cambridge 1963.
- R. Vernon, "International Investment and International Trade in the Product Cycle," *Quart. J. Econ.*, May 1966, 80, 190-207.
- H. Waehrer, "Wage Rates, Labor Skills, and United States Foreign Trade," in P. Kenen and R. Lawrence, eds., *The Open Economy: Essays on International Trade and Finance*, New York 1968.
- D. F. Wahl, "Capital and Labour Requirements for Canada's Foreign Trade," *Can. J. Econ.*, Aug. 1961, 27, 349-58.
- L. A. Weiser, "Changing Factor Requirements of United States Foreign Trade," *Rev. Econ. Statist.*, Aug. 1968, 50, 356-60.
- L. Weiss, "Concentration and Labor Earnings," Paper 6405, Social Systems Research Institute, University of Wisconsin, 1964.
- M. Yahr, "Human Capital and Factor Substitution in the CES Production Function," in P. Kenen and R. Lawrence, eds., *The Open Economy: Essays on International Trade and Finance*, New York 1968, 70-99.
- Organization for Economic Cooperation and Development (OECD), *Foreign Trade Statistical Bulletins, Series B, Commodity Trade*, Jan.-Dec. 1962, Paris 1963.
- U.S. Bureau of the Census, *Census of Manufactures, 1958, 1*, Washington 1961.
- , *Exports and Imports as Related to Output*, Series ES2, No. 5, Washington 1964.
- U.S. Bureau of Labor Statistics, *Wholesale Prices and Price Indexes, 1962*, Bull. 1411, Washington 1962.
- , *1/1000 Sample of the Population of the United States, 1960*, Social Systems Research Institute, University of Wisconsin, Madison.
- U.S. Department of Commerce, "The Transactions Table of the 1958 Input-Output Study and Revised Direct and Total Requirements Data," *Surv. Curr. Bus.*, Sept. 1965, 44, 45-49.
- U.S. Department of Interior, *1962 Minerals Yearbook, 2*, Washington 1963.

Optimal Restrictions on Foreign Trade and Investment

By FRANZ GEHRELS*

The theory of optimal restriction of international investment, and its interdependence with restriction on trade, has been developed in recent years by G. D. A. MacDougall, A. E. Jasay, M. C. Kemp, T. Negishi, and R. W. Jones (1967). The question introduced by MacDougall was whether, in analogy with trade interference, a borrowing country could alter the terms of lending to its advantage. Jasay considered the same question from the viewpoint of the capital-exporting country. Kemp and Jones went further to consider the interdependence between goods prices and factor prices in a two-country, two-good, two-factor situation; they found expressions for optimal duties and investment taxes, for both debtor and creditor countries. Of interest also is a recent note by V. K. Ramaswami on the related issue of taxing in a discriminatory way the earnings of immigrant labor. This shows that analogous rules can be developed for interference with the international movement of any factor of production—here we shall, however, restrict the discussion to the flow of goods and capital, keeping labor immobile between countries.

It is the purpose of this paper to modify the Kemp-Jones optimizing procedure, bringing it more nearly into line with

conventional tariff optimization. This permits easy extension to the n -goods case for both tariff and tax under full optimization (where both duty and tax can be adjusted freely). It also leads to more general answers for the cases of partial optimization dealt with by Jones (where either tax or duty can be varied, but not both). Jones took only the free trade case for the optimal investment tax, and only the free mobility of capital case for the optimal duty. A last problem we deal with is optimal restrictions on trade and borrowing for a country which has an imperfection in the domestic market for labor.

The method used here for finding optimal duties on trade, and taxes on income from foreign investment, is a natural extension of J. de V. Graaff's treatment of duties on several goods. That is to say, one maximizes a social utility function subject to a transformation function and a balance-of-trade constraint, by adjusting quantities of goods traded or the amount of capital invested. Marginal social utility of any good is assumed to match marginal private utility. Because of perfect competition and no externalities, consumer rates of substitution are equal to rates of transformation in production. At the maximum a Pareto optimum exists, because under competitive conditions and independence of tastes, each individual has maximized utility subject to his income constraint. There is no change which would increase any person's satisfaction without reducing that of others.

We make the conventional assumption that there is no foreign retaliation. The

* Professor of economics, Indiana University. This paper was completed during tenure of a research fellowship at the Institute for International Economic Studies, University of Stockholm, Sweden. I wish to acknowledge helpful comments on an earlier draft by the referee. He is absolved from responsibility for any remaining errors.

rest of the world may, however, have trade duties and investment taxes of its own. In that case the international terms of trade or lending facing the home country are different from those prevailing in the internal markets of foreign countries.

We wish to maximize a concave social utility function

$$(1) \quad U = U(C_1, C_2) = U(Y_1 - X_1, Y_2 - X_2)$$

where U_i will denote the partial derivative, C_1 and C_2 are consumption of good 1 and good 2, Y_1 and Y_2 are their outputs, and X_1 and X_2 are their exports.

The production constraint is written in the implicit form and takes account of the amount of capital invested abroad, and therefore not available for domestic use.

$$(2) \quad \phi(Y_1, Y_2, F) = 0$$

where ϕ_i denotes the partial derivative. The production possibility curve is concave to the origin in the (Y_1, Y_2) plane shifts downward whenever F , the amount of foreign investment, increases. The balance of trade is in equilibrium both before and after a transfer of capital and takes account of both commodities and invisibles. Following the other writers above, we do not concern ourselves with the adjustment process during the transfer. Therefore:

$$(3) \quad X_1 + X_2\pi + F\rho = 0$$

X_i positive represents an export and X_i negative, an import; π is the international terms of trade, with good 1 as numeraire; ρ is the foreign rate of interest, and F is the stock of external investment, in terms of good 1.

The free international mobility of capital leads to some complications with respect to the production possibility curve and the likelihood of incomplete specialization simultaneously for the home country and the rest of the world. We wish to make the production functions first-degree

homogeneous, and at the same time to have incomplete specialization everywhere a reasonable possibility. With only two factors of production, Jones (1967) showed that the probability of incomplete specialization both at home and abroad was zero.¹ In order to remedy this difficulty we as-

¹ With only two factors of production and constant returns to scale, complete specialization can occur rather easily under free mobility of either factor. This has been shown by Jones, *op.cit* (1967), pp. 31-38. When technologies are the same everywhere, the allocation of capital between countries and the outputs of goods in each are indeterminate. All we know from the Lerner-Samuelson factor-price theorem is that real wages and the rate of interest are equalized under incomplete specialization everywhere, whether or not capital is mobile. When conditions of production are *not* the same everywhere, the one-to-one mapping between goods prices and factor rewards is inconsistent with the free mobility of capital and production of *both* goods in each country. For example, let one country have a technology such that the given price implies a lower reward to capital than abroad: mobility of capital prevents the downward adjustment of the interest rate; unit cost is too high in the capital-intensive sector and so capital flows out until that sector disappears. The wage will now be determined in the one remaining sector by price *and* the supply of capital. The latter in turn is adjusted to the prevailing rate of interest, while the wage may be either higher or lower than abroad, depending on the technology.

Parenthetically, J. S. Chipman has, in a forthcoming paper, analyzed a special two-by-two case where incomplete specialization of both countries *can* happen, but where terms of trade and factor prices are invariant so long as both goods are produced in both countries. That is, the world production-possibility curve is flat on the segment corresponding to incomplete specialization. The conditions are that in one good the production functions be the same, between countries, while in the other the labor coefficients differ; in addition (this is not stated explicitly by Chipman) capital coefficients must differ in the opposite sense, so that unit cost with equal factor prices can be the same in both countries. These highly special conditions illustrate how difficult it is to obtain incomplete specialization in the two-by-two case. Moreover, the invariance of price and factor rentals makes the case uninteresting for the analysis of optimal interference with trade and lending. I am indebted to Chipman for letting me see his paper.

When a third factor, land, is present the unspecialized system is generally compatible with a fixed capital rental. In terms of Samuelson's equations and unknowns, we have six marginal productivity equations and two factor-supply equations to determine four independent factor ratios, two independent factor allocations, and the rewards to labor and land ($6+2=4+2+2$).

sume three factors of production; labor, capital, and land. It is then possible to have interferences with trade and foreign investment in addition to differences in technology, together with perfect international mobility of capital (or for that matter, labor), and still have both goods being produced in both parts of the world.

We must, however, pay a price for introducing a third factor of production, in terms of other complications. For one thing, the presence of land in both production functions means that labor's and capital's marginal productivities depend on two factor ratios rather than just one. In order still to say something definite about factor rentals, when duties on trade or taxes on foreign investment are adjusted, we shall assume that one good is strongly labor-intensive, and the other strongly capital-intensive. We take this to mean that a rise in price and output of the labor intensive good always makes labor scarcer and raises its wage; conversely, the rental on capital always goes up when the output of the capital-intensive good increases. A second complication is that, with one-product firms, the presence of the third factor may cause the market-guided goods-transformation locus to take peculiar shapes. We therefore assume explicitly that this path is everywhere concave to the origin.²

Taking account of the production and balance-of-trade constraints, we now re-

place the function U with the function W , where we use the trade constraint to eliminate X_1 , and we introduce the production constraint with a Lagrange multiplier.

$$(4) \quad W = U(Y_1 + X_2\pi + F\rho; Y_2 - X_2) \\ - \lambda\phi(Y_1, Y_2, F) = W(X_2, F)$$

We wish to maximize the function W with respect to both the volume of good 2 traded and the volume of capital invested abroad. Domestic outputs, Y_1 and Y_2 , and the international terms of trade and lending, π and ρ , are treated as dependent on both exports X_2 , and stock of foreign investment F .

A full optimum is obtained when X_2 and F can each be freely varied by policy makers, in which case the two are independent for purposes of differentiation. After finding full optima with respect to X_2 and F in the two following Sections I and II, we shall deal with the partial-optimum problem of Jones in Sections III and IV. In finding the optimal volume of trade, we treat F as dependent assuming there is some policy constraint on changing the degree of restriction on foreign investment. Similarly, in finding the optimal magnitude of F , we assume in Section IV that restrictions on trade are not freely adjustable, so that X_2 is dependent on F . Section V deals with a country having an imperfect domestic labor market. This situation we take to be representative of many developing countries.

I. Optimal Trade Duty

The objective of this section is to find the optimal tariff when policy makers have complete freedom to adjust restrictions on foreign investment, and in fact do maximize the social-utility function with respect to both variables. Holding F constant, we differentiate (4) with respect to X_2 and set the derivative equal to zero.

² When there are three factors and two goods, and any firm produces only one good, one can construct examples in which a factor's marginal product rises in the expanding industry and falls in the contracting one. This occurs because a factor's marginal product depends on its ratios to *both* other factors. Taking that factor's wage as numeraire, this means that marginal cost is falling in the expanding sector and rising in the other. Hence opportunity cost changes in the wrong way for concavity. Just as important, there is no longer a one to one mapping between goods prices and factor prices. For a fuller discussion see Gehrels. The claim there, however, was not quite correct: it is possible for the transformation curve to be convex everywhere to the origin.

$$\begin{aligned}
 \partial W / \partial X_2 &= U_1(\partial Y_1 / \partial X_2 + \pi \\
 &\quad + X_2 \partial \pi / \partial X_2 + F \partial \rho / \partial X_2) \\
 (5) \quad &+ U_2(\partial Y_2 / \partial X_2 - 1) \\
 &\quad - \lambda(\phi_1 \partial Y_1 / \partial X_2 \\
 &\quad + \phi_2 \partial Y_2 / \partial X_2) = 0
 \end{aligned}$$

Domestic outputs change whenever trade changes; and $\partial Y_2 / \partial X_2$ is always positive, whether good 2 is an export or an import, implying that $\partial Y_1 / \partial X_2$ must be negative. This follows from the condition $\phi_1 dY_1 + \phi_2 dY_2 = 0$, which is necessary for (2) to hold, given F constant, and that ϕ_1 and ϕ_2 are of the same sign. Thus dY_1 and dY_2 must be of opposite sign. The terms of lending depend on trade changes through adjustments in foreign production, given that the rest of the world is incompletely specialized. Which way ρ moves depends on which good is capital intensive.³

Because of perfect competition and my assumption of equality between social and private rates of substitution,

$$(6) \quad U_1 \partial Y_1 / \partial X_2 + U_2 \partial Y_2 / \partial X_2 = 0,$$

and

$$(7) \quad U_2 / U_1 = p,$$

where p is the domestic price ratio in terms of good 1.

Using (6) and (7), (5) now reduces to

$$\begin{aligned}
 (8) \quad &\pi + X_2 \partial \pi / \partial X_2 + F \frac{\partial \rho}{\partial X_2} - p \\
 &- \frac{\lambda}{U_1} \left(\frac{\phi_1}{\phi_2} + \frac{\partial Y_2}{\partial X_2} \bigg/ \frac{\partial Y_1}{\partial X_2} \right) \frac{\phi_2}{\partial Y_1 / \partial X_2}
 \end{aligned}$$

³ More generally, with constant returns to scale, but variable returns to proportion, the factor-price theorem of Samuelson states that (a) with an equal number of goods and factors there is a one-to-one mapping between goods prices and factor rentals, so that the terms of trade and lending must *both* move if either of them moves; (b) with more factors than goods (as is the case here with three factors and two goods, and a fortiori with complete specialization) there is a unique mapping from goods prices and factor supplies to factor rentals.

Since ϕ is the implicit form of the transformation function, ϕ_1 / ϕ_2 is the marginal rate of transformation between the two goods, i.e., the amount of good 2 obtained by giving up a unit of good 1 in production. But the expression in parentheses is zero, so that the ratio of production changes brought about by the change in trade is equal to the marginal rate of domestic transformation. The optimal tariff rate can now be written as

$$\begin{aligned}
 (9) \quad t &= \frac{\pi - p}{\pi} \\
 &\quad - \frac{X_2}{\pi} \frac{\partial \pi}{\partial X_2} - \frac{F}{\pi} \frac{\partial \rho}{\partial X_2} \\
 &= \frac{1}{\eta} \left[1 + \frac{F\rho}{X_2 \pi} \frac{E_\rho}{E_\pi} \right]
 \end{aligned}$$

which result agrees with those of Kemp and Jones.⁴ The coefficient η is the foreign price elasticity of demand for home-country exports of good 2 (or foreign supply elasticity, when good 2 is an import of the home country); $F\rho / X_2 \pi$ is the ratio of investment income to value of good 2 traded; and E_ρ / E_π is the elasticity of the interest rate with respect to the terms of trade. Through the "magnification effect" (see Jones 1965) this is generally greater than one. The reason for this is that a price is a weighted average of factor-rental changes; as output shifts, wages and the interest rate move in opposite directions, owing to our strong-intensity assumption above. On the other hand, $F\rho / X_2 \pi$ is generally less than one. The second term in the bracket could therefore be greater or less than one in absolute value, and have either sign. Its effect on the size of the

Thus, with given factor supplies there is again a connection between prices and rentals, and the latter must change with the former.

⁴ The equivalent of expression (9) has been found both by Kemp and by Jones. See Jones, (1967) equations (9) and (11) and pp. 8-16; also Kemp, equation (3) and pp. 792-95.

optimal duty (which could become a subsidy) therefore depends on particular assumptions.

To examine a particular case, let us assume that the home country is a net exporter of capital, and that its export good—let it be good 2—is capital intensive both at home and abroad. An increase in t , its trade duty, then increases π ; it also reduces the volume of trade, causes an increase in foreign production of good 2, and leads to a shift of home production away from good 2. Which way do the terms of lending move? From our assumptions they must move unambiguously in favor of the home country, because under full optimization, F the stock of foreign capital owned by the home country, is independent of the tariff rate. Thus the only fact of significance is that the rest of the world uses its capital more intensively (i.e., with more labor) in both employments than it did before.

One can easily see that a debtor country which also imports the capital-intensive good improves its terms of borrowing along with its terms of trade, when it raises its trade duty. This rather credible pair of cases suggests that the consequence of the terms-of-lending consideration for both creditor and debtor nations is the same: Each individually has an interest in restricting trade by more than the degree indicated by trade elasticity alone. Even if foreign demand elasticity were infinite, there may still be an argument for restricting trade—providing, of course, that there is no retaliation.

Without additional discussion it should be clear that when the home country is a capital exporter but exports the labor-intensive good, consideration of the return on foreign investment reduces the optimal duty on trade. Thus one can not make a general case either for higher or for lower duties on trade (before foreign retaliation) as a result of considering effects on the

terms of borrowing in addition to effects on the terms of trade.

Finally, it is natural to ask whether the situation is changed markedly when, contrary to the present assumption, the rest of the world is specialized completely. Our differentiating procedure was to hold F constant while differentiating partially with respect to X_2 . If the outside world is specialized in the numeraire good, then with a constant employment of all factors in that one good, the nominal return to capital abroad is independent of the terms of trade, and so of the duty. Hence the optimal tariff in (9) reduces to the reciprocal of price elasticity. But if the outside world is specialized in the non-numeraire good, then the rate of return on capital does depend on price. That is, the marginal physical productivity of capital stays constant as before, but its value productivity must change in the same proportion as price. (The elasticity of capital's return with respect to price is plus or minus unity.)⁵

We conclude this section by pointing out that it is possible to extend the procedure above without difficulty to that of n -goods produced and traded. Leaving its derivation to Appendix A, we obtain

$$(10) \quad t_k = \sum_{i=2}^n \frac{V_i}{V_k} \cdot \frac{1}{\eta_{ki}} - \frac{F\rho}{V_k} \cdot \frac{E_p}{E_{X_k}}$$

where η_{ki} is the foreign demand elasticity of the k th good traded with respect to the i th price, while V_i/V_k is the ratio of values in trade, of the i th and k th goods. The first term is therefore the weighted sum of reciprocals of price elasticities of the k th good with respect to all $n-1$ prices. The second term is the elasticity of the foreign

⁵ Kemp, Section III, does not distinguish between the two cases, even though he holds the foreign capital stock fixed. It is, however, unclear if he is referring to physical capital or value of capital in terms of the numeraire. Jones has the first half of the conclusion in the text above, but not the second half. See Jones (1967), Section III.

rate of interest with respect to the quantity of the k th good traded and weighted by the ratio of interest income to the value of the k th good in trade.

II. Optimal Investment Tax

In order to find the utility-maximizing tax on foreign investment, we now hold X_2 fixed and differentiate partially with respect to F . Domestic output is a function of F because a unit increase of capital invested abroad means a unit reduction of capital at home. This generally affects output of both goods, but differently according to their relative factor intensities.⁶ The terms of trade are also generally a function of F because goods prices are usually not independent of factor prices, and when the latter are affected by changes in F , the former are also. We obtain

$$(11) \quad \frac{\partial W}{\partial F} = U_1 \left(\frac{\partial Y_1}{\partial F} + X_2 \frac{\partial \pi}{\partial F} + \rho + F \frac{\partial \rho}{\partial F} \right) + U_2 \left(\frac{\partial Y_2}{\partial F} \right) - \lambda \left(\phi_1 \frac{\partial Y_1}{\partial F} + \phi_2 \frac{\partial Y_2}{\partial F} + \phi_F \right) = 0$$

From the side condition we obtain directly that

$$(12) \quad -r = -\frac{\phi_F}{\phi_1} = \frac{\partial Y_1}{\partial F} + \frac{\phi_2}{\phi_1} \frac{\partial Y_2}{\partial F} = \frac{\partial Y_1}{\partial F} + p \frac{\partial Y_2}{\partial F}$$

This states that the rate of interest, r , is equal to the marginal contribution of capital to national product, measured in

units of the numeraire good. This, in turn, is the negative of the sum of changes in output due to the transfer of a unit of capital abroad. Using, in addition, that $U_2/U_1 = p$, we reduce (11) to

$$(13) \quad X_2 \frac{\partial \pi}{\partial F} + \rho + F \frac{\partial \rho}{\partial F} + p \frac{\partial Y_2}{\partial F} + \frac{\partial Y_1}{\partial F} = 0,$$

and then to

$$(14) \quad \frac{\rho - r}{\rho} = -\frac{X_2}{\rho} \frac{\partial \pi}{\partial F} - \frac{F}{\rho} \frac{\partial \rho}{\partial F} = -\left(\frac{X_2}{\rho} \frac{\pi}{F} \frac{E_\pi}{E_\rho} + 1 \right) \frac{E_\rho}{E_F}$$

where E_π/E_ρ is the elasticity of foreign price with respect to the rate of interest. This is less than one, because price change is the weighted average of factor-rental changes, and, under our assumptions above, the wage rate falls when the interest rate rises. The ratio E_ρ/E_F is the elasticity of the foreign interest rate with respect to foreign investment by the home country. Let us again assume that good 2, the capital-intensive good, is exported by the lending country.

Some idea of the order of magnitude of the optimal tax on investment can be obtained by noting, as did MacDougall, that the elasticity of demand of a borrowing country for capital from abroad is the share of foreign-owned capital in its capital stock, times its total elasticity of demand for capital. The first term in the parentheses (14) can be either weaker or stronger than the second, because $X\pi/F\rho > 1$, being the ratio of visible to invisible exports in value terms, and $E_\pi/E_\rho < 1$. If their product were unity, the tax rate would therefore be about twice the elasticity of the foreign rate with respect to foreign investment.

⁶ Under the assumptions made for the factor-price equalization theorem, and in the two-good two-factor case, the Rybczynski theorem states that exporting a unit of capital will cause one output to fall and the other to increase. See T. Rybczynski. We can not invoke the theorem here because we have three, rather than two factors.

In the converse case, where the home country is capital-poor and thus both a net debtor and an importer of good 2, the optimal restriction of foreign borrowing is made greater by the fact that the foreign price of imported good 2 is reduced by the same action. In (9) X_2 and F from their definitions are both negative, while $\partial\pi/\partial F$ and $\partial\rho/\partial F$ remain negative; hence $\rho-r$ is now *negative*. The foreign return on capital is held below the return in the borrowing country.

We conclude that in the cases here described, the terms-of-trade and terms-of-lending effects reinforce each other for *both* creditor and debtor countries. Barring retaliation, optimal restrictions on trade and on lending will be more severe than when cross effects are neglected. It seems unnecessary to explore the case of X_2 being labor-intensive; but one can easily follow the same line of reasoning to show that the optimal restrictions on trade and lending would both be reduced in this case, and this is true for both creditor and debtor countries.

This result is related to that of Kemp and of Jones, if different in form. The reason for the difference lies in their maximizing social utility with respect to price, and to quantity of investment.⁷ Intuitively, what they have done can be viewed as follows: In the neighborhood of the optimum, transferring one unit of capital causes market disequilibria at initial prices both at home and abroad. Since foreign price can not be adjusted, equilibrium is restored by adjusting home-country prices, and so the tariff. By

⁷ This leads Jones to conclude that there can be a non-zero optimal tax on foreign investment only when there is a difference between domestic and foreign prices of goods, since under his assumptions neither price nor the interest rate in the unspecialized country is dependent on the stock of capital received from abroad. See Jones (1967) p. 12, and equation (12) and Kemp, p. 793, equation (2c), and p. 795, (5) and (6).

contrast, we have made quantity of exports independent, and allowed the secondary adjustment to the transfer of capital to take place in prices everywhere. Here too, some secondary adjustment of the tariff is needed.

For the n -goods case, leaving the derivation to Appendix B, we obtain the expression

$$(15) \quad \rho - r = - \sum_2^n X_1 \frac{\partial \pi_i}{\partial F} - F \frac{\partial \rho}{\partial F}$$

The excess of the foreign over the domestic rate of interest is thus the weighted response of the foreign interest rate, adjusted for all the price responses to the redistribution of the capital stock between countries. If, on balance, the home country exports capital-intensive goods, which are also capital-intensive abroad, then $\sum X_i \partial \pi_i / \partial F$ is negative, as is $\partial \rho / \partial F$. The previous conclusion, that optimal restrictions on lending are made more severe, is substantially unchanged.

III. Partial Optimization of Trade Duty

Governments are not always able to vary both the trade duty and the investment tax in order to achieve the kind of optimum defined above. It is therefore of interest to examine more generally the question considered by Jones, namely, how an optimal duty is obtained when the tax on investment cannot be varied.⁸ This section will show that the optimal duty subject to a given foreign-investment tax is higher than under full optimization, when the export good is also capital intensive. The corresponding modification of the optimal investment tax is left to the following section.

The main difference in finding an optimum lies in the fact that F , the amount of

⁸ Jones treats only the special case where the tax on investment is held at zero, so that $r=\rho$. See Jones (1967), pp. 23-31.

$$\begin{aligned}
 \frac{\partial W}{\partial X_2} = & U_1 \left(\frac{\partial Y_1}{\partial X_2} + \frac{\partial Y_1}{\partial F} \frac{\partial F}{\partial X_2} + \pi + X_2 \frac{\partial \pi}{\partial X_2} + F \frac{\partial \rho}{\partial X_2} + \rho \frac{\partial F}{\partial X_2} \right) \\
 (16) \quad & + U_2 \left(\frac{\partial Y_2}{\partial X_2} + \frac{\partial Y_2}{\partial F} \frac{\partial F}{\partial X_2} - 1 \right) \\
 & - \lambda \left[\phi_1 \left(\frac{\partial Y_1}{\partial X_2} + \frac{\partial Y_1}{\partial F} \frac{\partial F}{\partial X_2} \right) + \phi_2 \left(\frac{\partial Y_2}{\partial X_2} + \frac{\partial Y_2}{\partial F} \frac{\partial F}{\partial X_2} \right) + \phi_F \frac{\partial F}{\partial X_2} \right]
 \end{aligned}$$

foreign investment, is no longer independent of X_2 , the amount of exports. Differentiating (4) with respect to X_2 gives equation (16).

In order to reduce this expression, we have the following three relations:

$$(17) \quad \frac{\partial Y_1}{\partial X_2} + \rho \frac{\partial Y_2}{\partial X_2} = 0$$

$$(18) \quad \frac{\partial Y_1}{\partial X_2} + \frac{\phi_2}{\phi_1} \frac{\partial Y_2}{\partial X_2} = 0$$

$$(19) \quad \frac{\partial Y_1}{\partial F} \frac{\partial F}{\partial X_2} + \rho \frac{\partial Y_2}{\partial F} \frac{\partial F}{\partial X_2} = -r \frac{\partial F}{\partial X_2}$$

Since all three have been used in reductions above, they do not need further discussion. Equation (16) now becomes

$$\begin{aligned}
 (20) \quad \frac{\pi - \rho}{\pi} = & - \frac{(\rho - r)}{\pi} \frac{\partial F}{\partial X_2} \\
 & - \frac{X_2}{\pi} \frac{\partial \pi}{\partial X_2} - \frac{F}{\pi} \frac{\partial \rho}{\partial X_2}
 \end{aligned}$$

This expression differs in form from its counterpart (9) above only in taking account of any adjustment of the quantity of foreign investments to trade. In the special case where $\rho = r$ (the only one considered by Jones 1967), the expression reduces back to (9). However, this is not to say that the numerical result for the optimal duty is the same, because the values of the two remaining terms will differ according to the degree of interference with investment.

Given again that X_2 is exported by the home country and is capital intensive, and that there is incomplete specialization at home and abroad, an increase of the duty encourages foreign investment because foreign output of X_2 increases. Thus $\partial F / \partial X_2 < 0$, and $\partial \pi / \partial X_2 < 0$ just as before. If the foreign rental on capital is higher than that at home the gain from restricting trade is greater than in the independent case, and the optimal duty is higher.

Suppose now that the home country is instead the debtor, and that X_2 is the capital-intensive *import* good. Does the same expression show that a debtor country would restrict trade more severely than when it considers only the terms of trade and not the terms of lending? The answer is clearly yes because the signs of the three derivatives are the same, but $\rho < r$, $F < 0$, $X_2 < 0$, and $\pi < \rho$.

An intuitive interpretation of the extra term $\frac{(\rho - r)}{\pi} \frac{F}{X_2}$ is that when a lending country sends additional capital abroad because of a change in trade, and the foreign return is higher than the return to home, each unit gives it a net gain of $(\rho - r)$. This is additional to any induced change in the terms of lending. Conversely, when a debtor country obtains more foreign capital because of a change in trade, the gain for it is $(r - \rho)$ per unit, where this measures the excess of the domestic return over the foreign cost per unit of capital.

$$\begin{aligned}
 \frac{\partial W}{\partial F} = & U_1 \left(\frac{\partial Y_1}{\partial F} + \frac{\partial Y_1}{\partial X_2} \frac{\partial X_2}{\partial F} + X_2 \frac{\partial \pi}{\partial F} + \pi \frac{\partial X_2}{\partial F} + \rho + F \frac{\partial \rho}{\partial F} \right) \\
 (21) \quad & + U_2 \left(\frac{\partial Y_2}{\partial F} + \frac{\partial Y_2}{\partial X_2} \frac{\partial X_2}{\partial F} - \frac{\partial X_2}{\partial F} \right) \\
 & - \lambda \left[\phi_1 \left(\frac{\partial Y_1}{\partial F} + \frac{\partial Y_1}{\partial X_2} \frac{\partial X_2}{\partial F} \right) + \phi_2 \left(\frac{\partial Y_2}{\partial F} + \frac{\partial Y_2}{\partial X_2} \frac{\partial X_2}{\partial F} \right) + \phi_F \right] = 0
 \end{aligned}$$

IV. Partial Optimization of Investment Tax

We now assume that the tariff is fixed and that the tax on foreign investment is the policy variable. This problem was ingeniously analyzed by Jones but only for the particular case of free commodity trade (where $\pi = p$).⁹

We differentiate (4) with respect to F , to obtain equation (21). This time the side condition gives us that

$$\begin{aligned}
 (22) \quad r = \frac{\phi_F}{\phi_1} = & - \left(\frac{\partial Y_1}{\partial F} + \frac{\partial Y_1}{\partial X_2} \frac{\partial X_2}{\partial F} \right) \\
 & - \frac{\phi_2}{\phi_1} \left(\frac{\partial Y_2}{\partial F} + \frac{\partial Y_2}{\partial X_2} \frac{\partial X_2}{\partial F} \right)
 \end{aligned}$$

This again states that the loss of domestic output due to the transfer of a unit of capital matches the domestic rate of interest. After eliminations we get from (21) that

$$\begin{aligned}
 (23) \quad \rho - r = & - X_2 \frac{\partial \pi}{\partial F} - F \frac{\partial \rho}{\partial F} \\
 & - (\pi - p) \frac{\partial X_2}{\partial F}
 \end{aligned}$$

Let us first suppose that the home country is a creditor and an exporter of the capital-intensive good, X_2 . It is easily verified that

$$\frac{\partial X_2}{\partial F} < 0, \quad \frac{\partial \rho}{\partial F} < 0, \quad \frac{\partial \pi}{\partial F} < 0,$$

and $(\pi - p) > 0$ when there is a positive duty on trade. In expression (23), $(\rho - r)$ must therefore be positive because every term on the right is positive. The fact that exports of the capital-intensive good react negatively to capital exports, and thereby cause a loss per unit of export reduction measured by $(\pi - p)$, makes it worthwhile to restrict foreign investment by more than when trade is independent.

The expression works in an exactly symmetric way when the home country is a debtor and an importer of capital-intensive goods, because the derivatives are unchanged in sign, and $(\pi - p)$, F , and X_2 are now all negative. Restriction of foreign borrowing has an additional gain due to the increased importation of X_2 , weighted by $(p - \pi)$ per unit.

We conclude that, given the values of the price response and interest response terms in the optimal tax on investment equation (23), the optimal restriction on the international flow of capital is increased by the response of trade volume to investment. This conclusion however depended on the exportable good 2 being capital intensive. If good 2 were instead labor intensive the sign of the export response to foreign investment becomes positive ($X_2/F > 0$) and the extra term containing the price difference has the effect of reducing the optimal tax.

V. Optimal Interference for a Developing Country

The capital-poor country may in addi-

⁹ Jones (1967) pp. 15-23.

tion have distortions in the domestic market for factors of production, with the consequences for allocation discussed by Hagen, Bhagwati and Ramaswami, and Fishlow and David. Let the Y_2 sector (which is capital-intensive and import-competing) face a money wage rate higher than its opportunity cost in terms of labor's marginal product in the Y_1 sector. The consequence of this is (a) that Y_2 's money cost overstates its opportunity cost, so that its relative share in output is reduced from the perfectly competitive share; and (b), that the production-possibility curve is shifted inward because Y_1 uses too small a proportion of labor to other factors and Y_2 , too high a proportion. Marginal rates of transformation between factors are not equal between sectors. Allocation is inefficient on two counts and can be corrected fully only by correcting imperfections in the labor market. We wish to see how the rules above are modified when a labor-market distortion is introduced. That is, we seek a second-best optimum, subject to a given constraint in the labor market. The utility function to be maximized is the same, but the constraints are different. We have, as before,

$$(24) \quad \begin{aligned} U &= U(C_1, C_2) \\ &= U(Y_1 - X_1, Y_2 - X_2) \end{aligned}$$

The production constraint is now written as

$$(25) \quad \psi(Y_1, Y_2, F) = 0,$$

where, for a given domestic capital stock, the production-possibility curve lies inside that for ϕ (equation (2)) everywhere but at the end points, because of the labor-market imperfection. We have, in addition, that

$$(26) \quad U_2/U_1 = p = -b \frac{\partial Y_1}{\partial Y_2} = b \frac{\psi_2}{\psi_1}, \quad b > 1$$

This states that the marginal utility of

good 2 in terms of good 1 is greater than its opportunity cost in the constant ratio b . The necessary condition for a full optimum with respect to the duty becomes

$$(27) \quad \begin{aligned} \frac{\partial W}{\partial X_2} &= U_1 \left(\frac{\partial Y_1}{\partial X_2} + \pi + X_2 \frac{\partial \pi}{\partial X_2} \right. \\ &\quad \left. + F \frac{\partial \rho}{\partial X_2} \right) + U_2 \left(\frac{\partial Y_2}{\partial X_2} - 1 \right) \\ &\quad - \left(\frac{\partial \psi}{\partial Y_1} \frac{\partial Y_1}{\partial X_2} + \frac{\partial \psi}{\partial Y_2} \frac{\partial Y_2}{\partial X_2} \right) = 0 \end{aligned}$$

Using (26) this reduces to

$$(28) \quad \begin{aligned} U_1 \left[(1-b) \frac{\partial Y_1}{\partial X_2} + \pi + X_2 \frac{\partial \pi}{\partial X_2} \right. \\ \left. + F \frac{\partial \rho}{\partial X_2} \right] - U_2 = 0 \end{aligned}$$

$$(29) \quad \begin{aligned} \frac{p - \pi}{\pi} &= \frac{X_2}{\pi} \frac{\partial \pi}{\partial X_2} + \frac{F}{\pi} \frac{\partial \rho}{\partial X_2} \\ &\quad - \frac{(b-1)}{\pi} \frac{\partial Y_1}{\partial X_2} \end{aligned}$$

In examining this expression, note that now $X_2 < 0$ because X_2 is the import; $\partial \pi / \partial X_2 < 0$; $F < 0$ because the country is a net debtor; $\partial \rho / \partial X_2 < 0$ because good 2 is capital-intensive, and reducing its importation (dX_2 positive) causes the foreign interest rate to fall. Further, $\partial Y_1 / \partial X_2 < 0$, and $b-1 > 0$. Consequently, all three terms on the right-hand side are positive, and the effect of the labor-market imperfection is to *increase* the optimal rate of duty on imports. In common-sense terms, the overstated cost of good 2, and its consequent underproduction, is mitigated by the additional protection given to it. Even if the home country had no influence on the terms of trade or lending, it would still be worthwhile to restrict imports.

In order to find the optimal restriction on foreign borrowing, we hold X_2 constant and differentiate (24) with respect to F .

$$\begin{aligned}
 \frac{\partial W}{\partial F} = & U_1 \left(\frac{\partial Y_1}{\partial F} + X_2 \frac{\partial \Pi}{\partial F} \right. \\
 & \left. + \rho + F \frac{\partial \rho}{\partial F} \right) + U_2 \frac{\partial Y_2}{\partial F} \\
 (30) \quad & - \lambda \left(\psi_1 \frac{\partial Y_1}{\partial F} + \psi_2 \frac{\partial Y_2}{\partial F} + \psi_F \right)
 \end{aligned}$$

As before

$$\begin{aligned}
 (31) \quad -r = -\psi_F/\psi_1 = & \frac{\partial Y_1}{\partial F} + \frac{\psi_2}{\psi_1} \frac{\partial Y_2}{\partial F} \\
 = & \frac{\partial Y_1}{\partial F} + \frac{p}{b} \frac{\partial Y_2}{\partial F}
 \end{aligned}$$

$$(32) \quad r - p = X_2 \frac{\partial \Pi}{\partial F} + F \frac{\partial \rho}{\partial F} + p \frac{(b-1)}{b} \frac{\partial Y_2}{\partial F}$$

Because good 2 is capital-intensive, and importing more capital tends to reduce foreign production of good 2, and therefore to raise its foreign price, we have $\partial \pi / \partial F < 0$; importing more capital encourages home production of good 2, so that $\partial Y_2 / \partial F < 0$; and $\partial \rho / \partial F < 0$ because importing more capital tends to raise the foreign rate of interest. Remembering that X_2 and F are both negative, the terms of trade effect, if any, reinforces the terms of lending effect. But of greater interest is that the domestic cost distortion works counter to the first two effects. The first two terms are positive, but the third term is negative. The common sense of this is that the overstated costs of the import-competing sector, owing to excessive money wages, can be mitigated by reducing the home interest rate relatively to the international rate. If the home country has no influence on the international terms of trade or lending, expression (32) states that a developing country should *subsidize* foreign borrowing. That is, it should make the domestic rate of interest lower than the foreign rate.

VI. Summary

The general conclusion of this paper is

that the optimal restriction on trade or foreign investment is changed by the interrelation between goods prices and factor prices. Whether each restriction is increased or decreased thereby depends on the factor-intensities of the traded goods; but in general the two kinds of interference would be altered in the same direction.

When there is complete freedom to adjust both the duty rate and the tax on foreign investment, the optimal duty is increased for the exporter of capital-intensive goods if that country is also a creditor; the same statement is true for a country which exports labor-intensive goods and is a debtor. The opposite conclusions follow if factor intensities are switched.

The optimal tax on foreign investment or borrowing is increased under exactly the same conditions.

When the difference between international and domestic rate of return on capital is taken as fixed, the optimal duty is further modified because of induced transfers of capital. For the capital-intensive exporter and lender, the induced outflow of capital brings additional gain. For the labor-intensive exporter and borrower, the induced inflow brings an added advantage and implies a further increase of the duty.

When the international and domestic price ratios differ by a fixed proportion, the induced increase of trade makes it profitable for both creditor and debtor country to impose a larger tax on lending or borrowing than under 2.

When the debtor country also suffers from overstated costs in the import-competing, capital-intensive sector, because of differential wages, it becomes worthwhile to raise the trade duty further but to *reduce* the tax on foreign-owned capital.

APPENDIX

We shall here derive expressions (10) and

(15) in the text, which give the optimal tariff and optimal investment tax for the case of n goods.

A. Optimal Tariff

Expressions (1) through (4) of the text were the collective utility function, the production function in implicit form, the balance-of-payments constraint, and the constrained utility function. They now become

$$(1') \quad U = U(C_1, C_2, \dots, C_n) \\ = U(Y_1 - X_1, Y_2 - X_2, \dots, Y_n - X_n)$$

$$(2') \quad \phi(Y_1, Y_2, \dots, Y_n, F) = 0$$

$$(3') \quad X_1 + \sum_2^n X_i \pi_i + F\rho = 0$$

$$(4') \quad W = U\left(Y_1 + \sum_2^n X_i \pi_i + F\rho; \right. \\ \left. Y_2 - X_2, Y_3 - X_3, \dots, Y_n - X_n\right) \\ - \lambda \phi(Y_1, Y_2, \dots, Y_n, F) = 0$$

Just as before they describe the social utility function, the production function, and the balance-of-trade constraint. All but one of the X_i 's are independent. We have chosen to make X_1 a function of the other $n-1$ X_k 's, and have replaced it by means of the trade constraint (12). As in the two-goods case, all outputs are functions of each X_k , as are the $n-1$ international prices π_i , and the rate of return on foreign investment, ρ .

$$(5') \quad \frac{\partial W}{\partial X_k} = U_1 \left(\frac{\partial Y_1}{\partial X_k} + \pi_k + \sum_2^n X_i \frac{\partial \pi_i}{\partial X_k} \right. \\ \left. + F \frac{\partial \rho}{\partial X_k} \right) + U_{k-1} \frac{\partial Y_{k-1}}{\partial X_k} \\ + U_k \left(\frac{\partial Y_k}{\partial X_k} - 1 \right) + U_n \frac{\partial Y_n}{\partial X_k} \\ - \lambda \left(\phi_1 \frac{Y_1}{X_k} + \phi_k \frac{Y_k}{X_k} \right)$$

$$+ \phi_n \frac{Y_n}{X_k} \Big) = 0$$

$$k = 2, 3, \dots, n$$

In the same way as before, we have

$$(6') \quad \sum_2^n U_i \frac{\partial Y_i}{\partial X_k} + U_1 \frac{\partial Y_1}{\partial X_k} = 0,$$

and

$$(7') \quad U_i/U_1 = p_i = -\partial Y_1/\partial Y_i \\ i = 2, 3, \dots, n$$

With the use of (6') and (7'), (5') reduces to

$$(8') \quad \frac{\partial W}{\partial X_k} = \pi_k + \sum_2^n X_i \frac{\partial \pi_i}{\partial X_k} + F \frac{\partial \rho}{\partial X_k} \\ - p_k = 0$$

The expression for the optimal tariff is therefore

$$t_k = \frac{\pi_k - p_k}{\pi_k} = - \sum_2^n \frac{X_i}{X_k} \frac{\pi_i}{\pi_k} \frac{\partial \pi_i}{\partial X_k} \frac{X_k}{\pi_i} \\ - \frac{F\rho}{X_k \pi_k} \frac{E_\rho}{E_{X_k}}$$

Stated more simply, this is

$$(10') \quad t_k = \sum_2^n \frac{V_i}{V_k} \frac{1}{\pi_{ki}} - \frac{F\rho}{V_k} \frac{E_\rho}{E_{X_k}}$$

B. Optimal Investment Tax

For the n -goods case, we differentiate (4') with respect to F and obtain

$$\frac{\partial W}{\partial F} = U_1 \left(\frac{\partial Y_1}{\partial F} + \sum_2^n X_i \frac{\partial \pi_i}{\partial F} + \rho \right. \\ (11') \quad \left. + F \frac{\partial \rho}{\partial F} \right) + \sum_2^n U_i \frac{\partial Y_i}{\partial F} \\ - \lambda \left(\phi_1 \frac{\partial Y_1}{\partial F} + \sum_2^n \phi_i \frac{\partial Y_i}{\partial F} + \phi_F \right) = 0$$

We can simplify this by using the fact that

$$(12') \quad -r = -\frac{\phi_F}{\phi_1} = \frac{\partial Y_1}{\partial F} + \sum_2^n \frac{\phi_i}{\phi_1} \frac{\partial Y_i}{\partial F}$$

That is, the loss of home output due to a unit transfer of capital is the sum of all the output changes valued in terms of the numeraire good, and this in turn matches the rate of interest, under competition. Using (25) reduces the expression (24) to

$$(15) \quad \rho - r = - \sum_2^n X_i \frac{\partial \pi_i}{\partial F} - F \frac{\partial \rho}{\partial F}$$

REFERENCES

- J. Bhagwati and V. K. Ramaswami, "Domestic Distortions, Tariffs, and The Theory of Optimum Subsidy," *J. Polit. Econ.*, Feb. 1963, 71, 44-50.
- J. S. Chipman, "International Trade with Capital Mobility: A Substitution Theorem," in J. N. Bhagwati et al, eds., *Trade, Balance of Payments, and Growth; Essays in Honor of Charles Kindleberger*, Cambridge, Mass. 1971.
- A. Fishlow and P. David, "Optimal Resource Allocation in an Imperfect Market Setting," *J. Polit. Econ.*, Dec. 1961, 69, 529-46.
- F. Gehrels, "Factor Marginal Products and Decreasing Opportunity Cost," *Amer. Econ. Rev.*, Mar. 1965, 55, 114-21.
- J. de V. Graaff, "On Optimum Tariff Structures," *Rev. Econ. Stud.*, 1949, No. 1, 17, 47-59.
- E. Hagen, "An Economic Justification of Protectionism," *Quart. J. Econ.*, Nov. 1958, 72, 496-514.
- G. E. Jasay, "The Social Choice between Home and Overseas Investment," *Econ. J.*, Mar. 1960, 70, 105-13.
- R. W. Jones, "International Capital Movements and the Theory of Tariffs and Trade," *Quart. J. Econ.*, Feb. 1967, 81, 1-38.
- , "The Structure of Simple General Equilibrium Models," *J. Polit. Econ.*, Dec. 1965, 73, 557-72.
- M. C. Kemp, "The Gain from International Trade and Investment: A Neo-Heckscher-Ohlin Approach," *Amer. Econ. Rev.*, Sept. 1966, 56, 788-809.
- G. D. A. MacDougall, "The Benefits and Costs of Private Investment from Abroad: A Theoretical Approach," *Econ. Rec.*, Mar. 1960, 36, 13-35, reprinted in R. Caves and H. Johnson, eds., *A.E.A. Readings in International Economics*, Homewood 1968, 172-97.
- T. Negishi, "Foreign Investment and the Long-Run National Advantage," *Econ. Rec.*, Dec. 1965, 41, 628-31.
- V. K. Ramaswami, "International Factor Movement and The National Advantage," *Economica*, Aug. 1968, 35, 309-10.
- T. Rybczynski, "Factor Endowments and Relative Commodity Prices," *Economica*, Nov. 1955, 22, 336-41.
- P. A. Samuelson, "Prices of Factors and Goods in General Equilibrium," *Rev. Econ. Stud.*, Oct. 1953, 21, 1-20.

COMMUNICATIONS

Large Industrial Corporations and Asset Shares: Comment

By DAVID R. KAMERSCHEN*

In a recent issue of this *Review*, David Mermelstein, following the tradition of A. D. H. Kaplan and N. R. Collins and L. E. Preston, examined the changing shares of the 100 largest corporations, on a decade-by-decade basis over the period 1909-64. Mermelstein finds evidence of an increased ability in recent decades for the largest corporations to maintain their share of total assets. He then considers the "*progressive* or *growing* advantages that are possessed by these largest corporations" (p. 539). The purpose of this communication is to comment briefly on the alternative explanations offered by Mermelstein for the increased stability of the asset shares of the 100 largest corporations.

I think Mermelstein has made an important contribution to the empirical literature of industrial organization. To be sure, there are some serious data limitations in any study of the sort he has undertaken, but I do not feel this is an insuperable difficulty. Much more serious are the theoretical limitations underlying the "turnover" method. I do not think Mermelstein has satisfactorily answered the "hostile comments and reviews" (p. 531) that have been levied at the Kaplan turnover approach. Since most of these chastisements are cited in Mermelstein, I will not refer to them with the sole exception of quoting the following, and as far as I can see, yet unanswered criticism of George J. Stigler.

The statistical universe of the hundred or two hundred largest corporations is inappropriate to studies of monopoly

and competition, and we may hope that this [referring to Kaplan] will be the last study to fall prey to its dramatic irrelevance. For Kaplan's central idea—that the extent of instability in the relative fortunes of the leading firms is an informative symptom of competition—is important and deserves to be applied on a correct, industry basis. [1969, p. 338]

Although Mermelstein did not, of course, apply the Kaplan turnover method on "a correct industry basis," let me drop this general criticism in order to concentrate on his rationale for his empirical results. I have no serious questions concerning his first explanation, viz. the advantages of size.¹ Two minor objections that might be voiced are 1) his failure to cite what many would regard as probably the best empirical work that has been done on this topic by Marshall Hall and Leonard Weiss² and 2) his failure to tie in his argument more explicitly with the enormous rise in conglomeration in recent years (e.g., see Kamerschen 1970).

It is his second explanation that is my primary concern. This explanation is, he claims, a corollary of the basic Berle-Means-Lerner hypothesis of the separation of ownership and management.³ Mermelstein argues, fol-

¹ Stigler (1968) argues imperfections in the capital market is too often given as an explanation without any empirical evidence as to the transaction costs (see Kamerschen 1969b).

² Although a recent study by Richard Arnould cites both my work and Hall and Weiss as having demonstrated the "relationship between absolute size and the cost of capital to be highly significant and positive" (p. 74), this was a secondary theme in my paper and was not exploited with anywhere near the care and finesse with which Hall and Weiss operated.

³ I refer to the hypothesis in this way since Berle and Means first suggested the thesis, and Lerner recently documented that separation of ownership and management has in fact occurred in our economy over the last fifty years.

* Professor of economics, University of Missouri. I am grateful to Richard L. Wallace for his valuable comments and suggestions. The research reported here was financially assisted by a University of Missouri Summer Research Fellowship.

lowing R. Joseph Monsen and Anthony Downs, that if all firms were owner controlled there would be more variability in earnings and hence asset shares. "If instead all firms are managerial firms, then we would expect less variability in earnings and similarly fewer shifts in asset shares" (Mermelstein, p. 540).

In the end, whether owner and manager controlled firms perform differently is an empirical question. And it seems to me that the evidence is at best ambiguous with regard to the point Mermelstein is trying to make.

While Mermelstein does not cite any of the statistical findings on this point, at least four such studies have appeared. Three of these studies by Brian Hindley,⁴ Robert Larner, and Kamerschen (1968) provide no evidence to support the position that manager controlled firms perform differently in terms of profits than do owner controlled firms. The findings in these studies are now included in at least one introductory textbook. Thus Robert Lipsey and Peter Steiner (p. 370) state:

When Professor Larner sought to explain the significance of the difference between corporations with and without dominant ownership groups, he hypothesized that the manager-controlled companies should have lower profits and show less variation in profits if they were trying to avoid risks. The evidence when examined led him to reject the view of a significant difference in behavior.⁵

⁴ Hindley (1970) states that he has obtained results "similar to those reported by Kamerschen" using "different methods and data." At the time the present communication was written, I did not have access to this article.

⁵ They go on to say, "As is often the case, this led him to seek an explanation. He found that, although the members of the top management group need not be stockholders, they usually do hold sizable amounts of the stock in the corporation they manage, this stock often being acquired as a direct result of bonuses or compensation for their services. Most top managers of successful corporations are wealthy men, much of whose wealth is represented by stock in their own companies. For example, when Semon Knudsen left General Motors in 1968 to become President of Ford, he owned \$3.3 million of G.M. stock. In addition, Ford gave him 15,000 shares of its common stock worth \$850,000 and

Monsen, John Chiu, and D. E. Cooley did obtain results contrary to those of Hindley, Larner, and Kamerschen. That is, Monsen et al., found that the control status in the firms did influence profitability. While there has been at least one attempt to reconcile these differences (Kamerschen, 1969a) no definite conclusion can yet be drawn. As the author of one of the studies, I would like to conclude that the Hindley, Larner, and Kamerschen position is correct, but unfortunately, the empirical evidence at this time does not warrant such a conclusion. However, since my sample was larger and more comprehensive than that of Monsen and employed basically the same method of classifying firms as being owner or manager controlled, I think that Monsen and hence Mermelstein also must conclude that the issue is in doubt. While this in no way detracts from the overall value of Mermelstein's study, I would submit that his conclusions regarding the factors accounting for the largest corporations' *growing* advantages may have to be tempered in light of recent empirical evidence. I think this particularly applies in Mermelstein's case, for Larner's results indicated that the degree of control has no influence on either the *level* or the *variability* of profits, whereas Monsen found that the degree of control affected only the *level* of profits. And it is the *dispersion* rather than *level* that seems crucial in the Mermelstein study.

REFERENCES

- R. J. Arnould, "Conglomerate Growth and Profitability," in L. Garoian, ed., *Economics of Conglomerate Growth*, Corvallis 1969.
- A. A. Berle and G. C. Means, *The Modern Corporation and Private Property*, New York 1932; 2d ed. New York 1968.
- N. R. Collins and L. E. Preston, "The Size Structure of the Largest Industrial Firms,

an option to buy 75,000 more shares at a price below the price the public has to pay for it. Larner's study showed the total income of managers to be closely related to the profitability of the companies they managed" (pp. 370-71). Since the time the above passage was written, Knudsen has also left Ford.

- 1909-1958," *Amer. Econ. Rev.*, Dec. 1961, 51, 986-1011.
- M. Hall and L. Weiss, "Firm Size and Profitability," *Rev. Econ. Statist.*, Aug. 1967, 49, 319-31.
- B. Hindley, "Capitalism and the Corporation," *Economica*, Nov. 1969, 36, 426-38.
- , "Separation of Ownership and Control in the Modern Corporation," *J. Law Econ.*, Apr. 1970, 13, 185-221.
- D. R. Kamerschen, "The Influence of Ownership and Control on Profit Rates," *Amer. Econ. Rev.*, June 1968, 58, 432-477; correction Dec. 1968, 1376.
- , (1969a) "The Effect of Separation of Ownership and Control on the Performance of the Large Firm in the U. S. Economy," *Riv. Intern. di Scien. Econ. e Comm.*, July 1969, 61, 293-301.
- , (1969b) "Recurrent Objections to the Theory of Imperfect Competition," *Zeitschrift für die Gesamte Staatswissenschaft*, Oct. 1969, 125, 688-94.
- , "Conglomerate Mergers: The Myth and the Reality," *St. Johns Law Rev.*, Apr. 1970, 44, 133-51.
- A. D. H. Kaplan, *Big Enterprise in a Competitive System*, Washington 1954; rev. ed. Washington 1964.
- R. J. Lerner, "Ownership and Control in the 200 Largest Nonfinancial Corporations, 1929 and 1963," *Amer. Econ. Rev.*, Sept. 1966, 56, 777-87.
- , "Separation of Ownership and Control and its Implications for the Behavior of the Firm," unpublished doctoral dissertation, Univ. Wis. 1968.
- R. G. Lipsey and P. O. Steiner, *Economics*, New York 1966; 2d ed. New York 1969.
- D. Mermelstein, "Large Industrial Corporations and Asset Shares," *Amer. Econ. Rev.*, Sept. 1969, 59, 531-41.
- R. J. Mosen and A. Downs, "A Theory of Large Managerial Firms," *J. Polit. Econ.*, June 1965, 73, 221-36.
- R. J. Mosen, J. S. Chiu, and D. E. Cooley, "The Effect of Separation of Ownership and Control on the Performance of the Large Firms," *Quart. J. Econ.*, Aug. 1968, 72, 435-451.
- G. J. Stigler, "The Statistics of Monopoly and Merger," *J. Polit. Econ.*, Feb. 1956, 64, 33-40; reprinted in D. R. Kamerschen, ed., *Readings in Microeconomics*, Cleveland 1967; New York, 1969, 332-343.
- , *The Organization of Industry*. Homewood 1968.

Large Industrial Corporations and Asset Shares: Comment

By STANLEY E. BOYLE*

In a recent issue of this *Review*, David Mermelstein takes what, unfortunately, is not a very fresh look at the "share stabilization process on a decade-by-decade basis to find out whether the largest of these firms are better able to maintain their relative position during recent decades than those nearer the turn of the century" (p. 532).

While agreeing with the basic results obtained by Norman Collins and Lee Preston in 1961 and subsequently, Mermelstein argues that their work is marred by two serious shortcomings. First, the time periods employed by them (1909-19, 1919-29, 1929-35, 1935-48, and 1948-58) are of unequal length, making "meaningful comparisons difficult" (p. 532). Second, he believes that the approach they use "does not enable us to determine whether the type of stability increases or decreases the share of the very largest of the corporations being studied" (p. 532).

To offset the apparently overwhelming analytical problems raised by the Collins and Preston study, Mermelstein proposes to measure the changes in mobility exhibited by the 100 largest industrial corporations over the periods 1909-19, 1919-29, 1929-39, 1939-48, 1948-58, and 1958-64. Since some of the periods selected by Mermelstein are also of unequal length, the precise nature of this gain is unclear. Moreover, even if the periods selected were of equal length, it is questionable what may be gained from what is basically a repetition of the Collins and Preston "decade-by-decade" analysis. Analyses which involve such long periods ignore short-run changes in mobility and provide no insight into the timing and direction of changes which occur during a decade. As a consequence, any analysis of the effect of

merger activity or antitrust policy which is based upon such flimsy evidence is of debatable value. For these and other reasons, those portions of the paper are excluded from the comment.

Four specific questions are analyzed in this note: First, what is the analytical significance of mobility changes which abstract from changes in the composition of the group? Second, what is the basic relationship between long-run (decade) and short-run (two year) mobility changes which occurred over the period 1919-64? Third, what questions are raised regarding the economic significance of analyzing mobility changes for any larger specified number of firms (100 or 200)? Fourth, what is the nature of the problems inherent to any suppositions regarding the extent to which overall mobility and changes in the level of aggregate concentration may be related?¹

I. The Dropout Bias

"These [Mermelstein's] initial findings clearly suggest that increasingly the largest corporations are better able to hold onto their shares than their counterparts of an earlier era" (p. 534).

Mermelstein supports this basic conclusion with an analysis of changes in the rankings of *surviving* firms over a series of 10-year periods. The r^2 values obtained by Mermelstein compare the change in rank positions of these firms and are shown below as r_m^2 (all survivors only) and r_{ma}^2 (adjusted for industry shift). The values r_{bm}^2 which show the com-

* Professor of economics, Virginia Polytechnic Institute. I owe an immeasurable debt to my former colleague, Professor Joseph P. McKenna, who is responsible for many of the computations which are used throughout this paper.

¹ The data employed in this note includes *only* manufacturing corporations and excludes the mining and trade companies included in the earlier studies. The collection of the basic data employed in this and an earlier paper (Boyle and Sorensen) was begun while the author served as Chief, Division of Industry Analysis, Federal Trade Commission. These differences should not affect the observations included in this comment since they are restricted to the method employed by Mermelstein rather than the actual statistical results obtained in his study.

TABLE 1—CHANGES IN ASSET-SIZE RANKINGS FOR ALL SURVIVORS (ADJUSTED AND UNADJUSTED) AND ALL FIRMS: 1909–58

Years Compared	r_m^2	r_{ma}^2	r_{bm}^2
1909, 1919	0.92285	.81045	—
1919, 1929	0.73704	.88654	0.407
1929, 1939	0.91929	.94472	0.678
1939, 1948	0.92686	.95555	0.691 ^a
1948, 1958	0.93864	.93968	0.649 ^b

^a For the period 1939–49.

^b For the period 1949–59.

Source: The r_m^2 and r_{ma}^2 value are taken from Mermelstein's *Review* article. The r_{bm}^2 are taken from Boyle and McKenna's forthcoming article, Table 1.

parable results obtained by Stanley Boyle and Joseph McKenna are computed on an "all firms" rather than an all survivors basis.²

It is interesting to note that the r_{ma}^2 values in Table 1 follow roughly the same path as do the r_{bm}^2 values. The most prominent difference between them is that they show substantially different levels. In both cases, however, the r^2 values are lowest for the 1919, 1929 comparison and then rise over the next two periods declining slightly in the last. The difference in level is, however, an exceedingly important property of the comparison. If the fact that there has been relatively little movement in rank of leading firms is economically significant, and presumably related to the structure of the industry, then the size of the segment of industry under examination is of considerable importance. Put another way, the decade of the 1920's was one of considerable mobility because 1) the rank of surviving firms changed, and 2) new firms entered and old firms departed from the charmed circle (the 100 largest).³ This latter movement indicates that the rise of new competitors may well be more important than internal rank changes.

² In addition to the differences in coverage, i.e., all survivors compared with all firms, the Boyle-McKenna data begin in 1919 rather than 1909, and actually pertain to decades, i.e., 1939–49 and 1949–59 rather than 1939–48 and 1948–58.

³ Mermelstein does make an interesting contribution here in attempting to measure the importance of industrial shift upon rank stability. This was attempted in a much less precise matter by Boyle and McKenna for the period 1919–64 (pp. 10–14).

The r_m^2 and r_{ma}^2 values for the period 1919–29 are .737 and .887, respectively, compared with the r_{bm}^2 value of .407. The bulk of this difference is due to the fact that the latter value takes explicit account of the fact that by 1929 only 72 firms remained of the 100 largest of 1919. Carried to its most absurd extreme the "survivors only" approach for analyzing industry stability would produce an r_m^2 value of 1.00 if the leading firm alone among the top 100 retained its position of the period in question *but* the identity of all of firms ranked 2–100 changed over the period. Obviously, a value for $r_m^2 = 1.00$ in that situation is without economic significance. Therefore, the failure to note this second aspect of mobility, i.e., in and out of the group, and taking into account only the changes of those included in *both* periods constitutes a basic and substantial analytical shortcoming of the Mermelstein analysis.

II. The Decade Problem

As important as the preceding problem is, its significance for competition is substantially outweighed by the fact that the measurement of mobility over a period as long as a decade ignores almost entirely the timing and direction of changes in large-firm stability. The differences between the paired-year and decade stability coefficients for the period 1919–39 are shown in Table 2.⁴

As would be expected, the data show greater stability (higher r^2 values) over the relatively short two-year periods than they do over the decade. More importantly, they show that mobility was the greatest between 1919 and 1923. Between 1923 and 1929 it remained almost constant, increased briefly between 1929 and 1931 and then declined sharply after 1931. Thus, examining only the decade movement obscures both the magnitude and timing of the significant movements which actually occurred between 1919 and 1931.

Looking at the paired-year values shown in Table 2 for the postwar period, it is possible to follow the progress of the large merger movement. Ignoring the period prior

⁴ The data for the entire period (1919–41 and 1948–64) are included in Table 2.

TABLE 2—PAIRED-YEAR AND DECADE STABILITY COEFFICIENTS: 1919, 1921 to 1962, 1964

Years Compared	Paired-Year Stability Coefficients	Decade Stability Coefficients	Time Period	Paired-Year Stability Coefficients
1919, 1921	.763		1939, 1941	.840
1921, 1923	.738			
1923, 1925	.800	.407	1948, 1950	.916
1925, 1927	.805		1950, 1952	.879
1927, 1929	.802		1952, 1954	.898
1929, 1931	.770		1954, 1956	.831
1931, 1933	.871		1956, 1958	.902
1933, 1935	.931	.678	1958, 1960	.859
1935, 1937	.886		1960, 1962	.845
1937, 1939	.907		1962, 1964	.841

Source: See Boyle and McKenna; Table 1.

to 1948 which was affected strongly by the economic changes growing out of World War II, we see that the highest degree of stability was shown between 1948 and 1950. Since then, however, with minor exception *apparent* mobility has increased.⁵ The decade data alone fail to describe this pattern of steady change. Thus, the faith evidenced in, and the analytical conclusion drawn from decade averages such as those presented by Mermelstein must be tempered in light of their obvious shortcomings. Long-term stability data are useful *only* if proper attention is devoted the concomitant short-term changes which occur within the longer period.

III. The Relevant Size Class

A third serious, but scarcely unique problem of the Mermelstein paper surrounds its preoccupation with the changes which occur among the 100 largest industrial concerns. While the use of this level of firm aggregation may be of some value as an interim analytical device, one should remember that data for that group have no unique economic significance. There is little, if any, evidence that

data for *any* particular number of firms has *any* critical economic value, as George Stigler pointed out in his critique of the Kaplan mobility findings more than a decade ago.

It is obvious that few, if any, economists, Stigler among them (1966, pp. 232-33), have paid sufficient attention to this admonition in recent years. The advice is, nonetheless, sound. Never put all of your eggs in one basket, particularly an irrelevant one: Two of the studies referred to above (see Boyle and McKenna, Boyle and Robert Sorenson) do analyze other levels. The second examines intra-industry mobility at the 10, 20, and 50 largest firm levels. Both of these papers show that two significant changes occur as the number of firms under examination expands. First, the level of the r^2 values tend to increase. Second, they tend to become somewhat more stable, i.e., they exhibit essentially smaller shifts between years.

In an effort to portray these changes over the period 1919 and 1964, the 100 largest firms have been subdivided into four groups (1-25, 26-50, 51-75, and 76-100) and paired-year stability coefficients computed for them. Table 3, shows the mean \bar{x} and standard deviation σ of the paired year stability coefficients for each asset-size sub-group over each of the four major subperiods (1919, 1929; 1929, 1939; 1948, 1956; and 1956, 1964).⁶

The data presented in Tables 2 and 3 show that while the level of the stability coefficients for the 100 largest industrial concerns declined somewhat since 1948, that was not true for the 25 largest. On the contrary, the mean (\bar{x}) of stability coefficients for the subgroup of largest firms actually increased in each of the subperiods. At the same time the σ values declined. Clearly mobility declined for firms in the top size class. Movements in the other size classes show lower and more variable \bar{x} values and larger values for σ . In some years the \bar{x} values for the 76-100 size class group are zero or even negative. They show no significant

⁵ Recent discussion of this phenomenon (Boyle and Sorenson) shows that "real" mobility (that which is not a function of changes wrought by merger activity) actually declined between 1950 and 1964 in the seven 2-digit industries which were examined.

⁶ The paired-year coefficients for each of the component size classes are available from the author by writing department of economics, V.P.I., Blacksburg, Virginia 24061.

TABLE 3—COMPUTED MEAN \bar{x} AND STANDARD DEVIATION σ OF PAIRED-YEAR STABILITY COEFFICIENTS BY SIZE CLASS

Size Class	1919, 1929 ^a		1929, 1939 ^a		1948, 1956 ^b		1956, 1964 ^b	
	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
1-25	.6544	.2445	.8989	.2068	.9520	.0161	.9622	.0067
26-50	.5567	.3053	.4820	.3023	.7550	.0502	.4569	.2344
51-75	.4324	.0824	.6886	.1847	.6106	.2289	.6007	.1414
76-100	.4497	.2126	.3164	.2874	.4522	.1263	.5607	.1707

^a The values shown are based upon paired-year coefficients covering 5 time periods.

^b The values shown are based upon paired-year coefficients covering 4 time periods.

pattern. On the average they seem to show smaller values for \bar{x} but higher σ values.

The basic reason for this divergence in the \bar{x} and σ value is the extent to which merger activity has affected the membership of each group. The approach used by Mermelstein to determine the impact of mergers (pp. 536-38) fails to take adequate account of the impact of the merger movement of the past decade and one-half.⁷ As Boyle and McKenna pointed out: "An outstanding example of stability appears in the very largest firms. The 12 largest manufacturing firms in 1929 remained among the 12 largest firms in 1962, one-third of a century later" (p. 5).

While Mermelstein and others may have focused their analysis only upon changes among the 100 or 200 largest industrial corporations, and while these may be convenient numbers to use for the sake of presenting consistent estimates of the control over economic resources of some fixed number of firms through time, it should *not* be inferred that such estimates necessarily contain specific economic significance.

IV. Stability and Concentration

Although Mermelstein does not specifically direct his attention to the relationship between the level of aggregate concentration and the degree of stability which exists, it seems worthwhile to raise this question briefly at this time. Unfortunately, it is not one to which precise answers can be given. However, the data used in this comment

taken in conjunction with those which appear in a recent Federal Trade Commission Report show an interesting relationship. The FTC study shows that the share of total corporate manufacturing assets held by the 100 largest firms increased from 39.4 percent in 1951 to 47.1 percent in 1958, holding at about that level until 1966. In 1968 and 1969 it rose again (p. 176).

Table 4 may serve to indicate the relationship between changes in stability and changes in the level of aggregate concentration for subgroups within the 100 largest over the period from 1950 to 1962. These data show that the increase in the level of concentration among the 100 largest corporations (5.5 percentage points) was the summary result of growth by the 10 largest firms (FTC p. 121). These data, taken in conjunction with those shown earlier (Boyle and McKenna), indicate strongly that when appropriate levels of aggregation are selected there may indeed be a strong correlation

TABLE 4—CHANGES IN CONCENTRATION, BY ASSET-SIZE CLASS WITHIN THE 100 LARGEST CORPORATIONS: 1950-1962

Size Class	Change in Concentration	Cummulative Change
	(percentage points of change)	
1-10	3.1	3.1
11-20	0.2	3.2
21-50	0.9	4.1
51-100	1.4	5.5

Source: Computed from, Willard Mueller, "Statement," p. 121. The figures used represent slight differences from those contained in the original table and are based upon a revision of those data.

⁷ The reader interested in this area is directed to: FTC, Economic Report on Corporate Merger, particularly chapters 4 and 5.

between stability, i.e., the ability of old firms to maintain their relative position, and changes in concentration.

V. Conclusions

Although my comments have been numerous and rather pointed, it should not be inferred that Mermelstein's paper is without merit. Rather they should be interpreted as an attempt to indicate the type of problems which may ensue when economists and others are too long bound to the use of limited data sets which tend to obscure rather than explain significant changes through time.

REFERENCES

- S. E. Boyle and J. P. McKenna, "Size Mobility of the 100 and 200 Largest U.S. Manufacturing Corporations. 1919-1964," *Anti-trust Bull.*, forthcoming.
- S. E. Boyle and R. L. Sorensen, "Concentration and Mobility: An Alternative Measure of Industry Structure," *J. Ind. Econ.*, forthcoming.
- N. R. Collins and L. E. Preston, "The Size Structure of the Largest Industrial Firms, 1909-1958," *Amer. Econ. Rev.*, Dec. 1961, 51, 986-1011.
- A. D. H. Kaplan, *Big Enterprise in a Competition System*, rev. ed., Washington 1964.
- D. Mermelstein, "Large Industrial Corporations and Asset Shares," *Amer. Econ. Rev.*, Sept. 1969, 59, 531-41.
- W. F. Mueller, "Statement," *Economic Concentration: Part I*, Hearings before the Senate Subcommittee on Antitrust and Monopoly, Washington 1964, 109-29.
- L. E. Preston, "Statement," *Economic Concentration: Part I*, Hearings before the Senate Subcommittee on Antitrust and Monopoly, Washington 1964, 56-70.
- G. J. Stigler, "The Statistics of Monopoly and Merger," *J. Polit. Econ.*, Feb. 1956 64, 33-40.
- , "The Economic Effect of the Anti-trust Laws," *J. Law Econ.*, Oct. 1966, 9, 225-37.
- U.S. Federal Trade Commission, *Economic Report on Corporate Merger*, October 1969.

Large Industrial Corporations and Asset Shares: Reply

By DAVID MERMELSTEIN*

By bringing to my attention a body of data I evidently overlooked, David Kamerschen properly disputes my suggestion that differential patterns in managerial behavior—patterns relating to corporate control—may partially explain the increased ability during recent decades of the largest corporations to maintain their share of total assets.¹ Since Kamerschen does not challenge what I consider to be the major contribution of my article—its methodology and empirical findings—we are left, even more than before, with a shortage of possible valid explanations for the upward trend in the regression coefficients.

One possible explanation of my empirical findings, not mentioned in my previous article, may be found in the activities of the state. For example, substantial indirect subsidies have been given to the oil and automobile industries through construction, maintenance, and administration of a vast network of roads and highways.² Subsidies do not exist solely on the expenditure side of the state budgetary ledger. By permitting individuals to deduct mortgage interest payments from their taxable income, as well as state and local property taxes, and by not requiring individuals to report the imputed rental value of owner occupied homes, the state not only contributes to urban and sub-

urban sprawl but decentralizes the population in such a way as to increase the demand for automobiles and hence indirectly subsidizes this industry.³

This argument has obvious merit in explaining my empirical results. It is not, however, without its limitations. Should all firms in an industry share in government subsidies equally (that is, in proportion to their size), then no increase should take place in those regression coefficients which have been adjusted for changes in industrial structure. To the extent that the largest firms receive a greater than average share of the complex mix of direct and indirect subsidies—a highly plausible hypothesis, to say the least—then some increase would occur in the regression coefficients over time, even when adjustments are made for changes in industrial structure.⁴

Turning next to the comment of Stanley Boyle, I am gratified that his research tends, by and large, to confirm my own. As Boyle himself points out, the values of r_{ma}^2 and r_{bm}^2 (in his Table 1) "follow roughly the same path." Moreover, his disaggregation of the data, presented in Tables 2-4, seems to be consistent with the work I have done and more in the nature of a complement than a substitute.

On the other hand, in the process of consistently presenting my research as though it were an analysis of changes in corporate *ranks* as well as in focussing on my parenthetically mentioned measuring rod of stability, r^2 the coefficient of determination, Boyle has unfortunately misrepresented my article and its somewhat different approach to industrial stability. Ranks, per se, were

* Polytechnic Institute of Brooklyn.

¹ Kamerschen is less than fair, however, when he claims that I failed to tie in my explanations "more explicitly" with the "enormous rise in conglomeration in recent years." I twice mention the effects of a recent increase in diversification and in the second of these discussions considerable space is allotted to this problem. In fn. 19 of my original article, I make it clear that in referring to diversification, I had conglomerateness largely in mind.

² A disproportionate number of the twenty largest firms have been in the oil and automobile industries since 1909:

1909	1	1929	10	1948	10	1964	10
1919	7	1939	10	1958	11		

Source: Mermelstein (1967, pp. 90, 97-150).

³ For a brilliant and penetrating analysis of state expenditures in a capitalist society, see J. O'Connor (1970).

⁴ The interested reader is referred to my dissertation for a discussion of other factors, such as the market for managers, bureaucratic practices, and dividend disbursements that may help to explain the upward trend of the regression coefficients.

not even studied and in analysis of what I did study—the relative shares of corporate assets—use was made of the regression coefficient, which in contrast to r^2 , as explained in the original text, measures a different kind of stability from the latter. It is perfectly possible, for example, for r^2 to be higher during recent decades (or during Boyle's paired-years) with b yet remaining constant. This would indicate no increase, on average, in the ability of the largest firms to maintain their shares, but rather some decrease (as indicated by the higher r^2) in erratic changes in shares (or ranks).

Because he does not notice the differences between an analysis based on comparisons of regression coefficients and one based on coefficients of determination, Boyle errs when he asserts that my "survivors only" method has a "basic and substantial shortcoming." My attempts to check out the biases involved in this procedure show it to be, most probably, a reliable and unbiased method, *given the uses to which it was put*. Boyle is of course correct, though hardly relevant, in pointing out that a single firm would have an r^2 of unity. In actuality, the number of survivors for the later periods stabilized in the mid-1930's, and even the number in the earliest and most volatile decade remained relatively high at 61: 1909-1919, 61; 1919-1929, 70; 1929-1939, 85; 1939-1948, 84; 1948-1958, 85; and 1958-1964, 88. A more extended argument that the "survivors only" method is of negligible bias, I relegate to the Appendix.

Elsewhere in Boyle's comment, and in Kamerschen's as well, a question is raised concerning the relevance of certain kinds of research engaged in by students of industrial organization. Both (independently) approve George Stigler's statement that "The statistical universe of the hundred largest corporations is inappropriate to studies of monopoly and competition, and we may hope that this (referring to Kaplan) will be the last study to fall prey to its dramatic irrelevance."⁵ Research on "the extent of

instability in the relative fortunes of the leading firms," Stigler continues, should be "applied on a correct, industry basis." (In passing, if Stigler is correct, then Kamerschen's statement that "... Mermelstein has made an important contribution to the empirical literature of industrial organization" is gratuitous.)

Stigler's viewpoint, as expressed in the passage quoted by Boyle and Kamerschen, is of considerable importance.⁶ Behind it lie key assumptions about the nature of the capitalist process and the relationship between economic phenomena and the political institutions within which economic decision making takes place in this country. Many economists share Stigler's sentiments; nonetheless, I believe he is mistaken. Without, I hope, violating the spatial privileges given an author to reply, I would like to set down a few introductory comments suggesting an alternative framework in which these issues can be judged.⁷

To begin, it is not self-evident that conglomerateness, and by the same token, the statistical universe of the one or two hundred largest industrial corporations, has no relation whatsoever to economic power in the marketplace. For example, Corwin Edwards in a statement to a Senate Subcommittee lists four ways in which conglomeration enhances market power: a) subsidization; b) reciprocity; c) full line selling; and d) the forbearance that prevails among large conglomerates (pp. 43-45). Mere size alone may

factorially answered the 'hostile comments and reviews' " [of Kaplan's book]. To the contrary, my own approach to this aspect of industrial stability uses what I consider a more refined statistical *alternative* to the turnover method used by Kaplan. Kamerschen indicts my study by declaring it to have "theoretical limitations." Let us be clear that these charges do not apply to the area of method at all, either Kaplan's turnover process or my regression analysis, but to the subject matter itself—the universe of the largest corporations.

⁶ Another leading spokesman of this point of view, Morris Adelman, in a statement before a Senate Subcommittee, has asserted that "absolute size is absolutely irrelevant" and by implication that neither conglomeration nor increased overall concentration can ever have any harmful competitive effects (p. 228).

⁷ Many of these comments are developed at greater length in the introductory material and a number of the readings in Mermelstein (1970).

⁵ In the interests of clarity, it should be pointed out that Kamerschen errs in implying that I, like A. D. H. Kaplan, use the "turnover" method and have not "satis-

also be instrumental in obtaining less expensive credit⁸ or in perpetuating industrial control of various markets through the ability to purchase the pertinent technology. More important, absolute size has always been thought to be crucial in times of war. This time honored precept applies not merely to nations but to industrial corporations as well. Nor need wars break out for power to be exercised: threat of war by itself is often sufficient.

It is interesting to ponder the effects on pricing, output, resource allocation, and technological development if there existed a total of only one hundred firms in the entire economy, each firm of equal size, and each operating in every existing industrial market. With (4-firm) concentration ratios of 4 percent, this economy, by Stigler's judgment, must be super-competitive. Remembering Edwards' discussion, it would be rash to assert that there would be absolutely no difference between this situation and one in which each industry contained one hundred equal-sized firms, not one of which operated in more than one industry.

To pursue the matter one step further, have we collected industrial data and fashioned our definitions of industry with such theoretical precision that it can be said with perfect assurance that studies of overall concentration have only "dramatic irrelevance?" The current merger movement⁹ contains an extraordinary number of firms desperately striving to avoid dependence on a single industrial market. John Galbraith's *The New Industrial State*, whatever its failings,¹⁰ correctly perceives that a substantively new pattern of industrial organization has been created, or at least is in the process of being developed, dissimilar only in degree from my hypothetical example above (where total

output was produced by one hundred firms). What is happening sorely needs further study, not casual dismissal out of hand.

The primary reason for concern is that power in the market place is but one facet of industrial concentration. Power also exists to influence or control key governmental decisions; or, in other words, industrial concentration has a political dimension of vast importance.¹¹ It is no accident, for example, that Presidents have been prone to appoint bankers, businessmen, and Wall Street lawyers to cabinet posts and other policy-making positions inside the executive branch of government.¹² This power is continually exerted in such areas as foreign policy, military contracts, labor, taxation, budgetary programs, urban policies, transportation, and basic research.

As an example, take two problems very much in the news these days—air pollution and urban mass transit. No one can question their importance: the condition of the latter is part of the breakdown in our urban centers while the former is beginning to threaten the life process itself.¹³ Yet nothing much is done. This should come as no surprise since lasting solutions to these and other problems of American society require radical changes of vast magnitude. At the very least, we need an all-out attack on the population problem, rejection of economic growth as a way of life,¹⁴ and comprehensive regional planning involving the location of jobs, housing, and

¹¹ For a classic discussion of some of the political aspects of oligopolistic rivalry, see K. W. Rothschild. As Rothschild points out, "The fact is that when we enter the field of rivalry between oligopolistic giants, the traditional separation of the political from the economic can no longer be maintained. . . . *The oligopolistic struggle for position and security includes political action of all sorts right up to imperialism. The inclusion of these 'non-economic' elements is essential for a full explanation of oligopoly behavior and price*" (pp. 462-63, as reprinted in *Readings in Price Theory*; italics in original).

¹² Evidence that American society is ruled by an upper class based on possession of wealth is contained in G. W. Domhoff.

¹³ For a sober appraisal, see B. Commoner.

¹⁴ What follows, by author Edward Abbey (quoted by Chapman and Harrington), is the best succinct statement I have come across expressing by analogy the crux of the problem: "Growth for the sake of growth is the ideology of the cancer cell."

⁸ Donald Dewey's argument (quoted in my original article), that advantages in the capital market may be due to the government aid that giant firms are likely to receive when threatened by bankruptcy, is apropos in this context.

⁹ For a good discussion, see Reid pp. 73-120.

¹⁰ For example, Galbraith shows little interest in the overseas activities of the American corporations and in the rise of the multi-national corporation. See also R. Miliband, reprinted in Mermelstein (1970).

recreation in which the real needs¹⁵ of the people are serviced.¹⁶

Unfortunately, these changes are in sharp conflict with existing property rights and the way in which corporations have defined their economic interest as well as their preferred modes of operation. Consequently, pollution, congestion, deteriorating mass transit, as well as the other social ills we associate with urban blight, do not just happen. They are the end products of a chain of decisions the first links of which have been forged by those at the apex of the corporate system. Increasingly, economic growth—the paramount goal of American life¹⁷—has been spearheaded by the largest of these corporations. Their decisions have structured American society. Their vision is one of continued expansion of one kind of consumption—goods sold in the marketplace for profit—giving short shrift not only to the alternatives of public goods and greater leisure but to the social and ecological repercussions as well. This vision has molded the consciousness of the American people, and while the corporations give lip service to the ideology of the free market, they do not

hesitate to use political means to perpetuate their ideological and economic hegemony.

The crisis in pollution and transportation is not unique. Whether the issue is crime or war, miseducation or addiction, poverty or racism, solutions are nowhere in sight. Instead, we have a vast network of government subsidies, government controls, and government programs of all kinds, many in conflict with each other, but virtually all in service of one or another industrial interest.

Corporate capital has never been so powerful as it is today,¹⁸ given the relative displacement of other propertied classes such as the independent farmer and the small businessman.¹⁹ Increasing concentration of assets, more secure than ever, serves to ease the task of coordinating decisions in defense of corporate hegemony, a task of some importance for those who occupy the upper reaches of the corporate world, given the current situation in America, one in which growing numbers of Americans, especially its youth, are deeply alienated and openly hostile to the capitalist way of life.

APPENDIX

In the September 1969 issue of this *Review*, I used the following regression procedure to determine mobility of large corporations. For each time period, the share of survivor assets is regressed on their corresponding early year shares. A regression coefficient greater than unity means

¹⁵ Without minimizing the immense difficulties involved in defining this term, I would suggest a meaningful approach is to be found in the social and biological requirements needed to preserve the species: nutrition, protection from environmental hazards, reproduction, and emotional well-being. (I am indebted to my colleague, Professor Shane Mage, for this approach. Helpful insights about the meaning of human needs have also come from a variety of sources, especially Galbraith, P. Baran, P. Baran and P. Sweezy, and H. Marcuse.)

¹⁶ Redirection of social priorities and productive energies may be the beginning of the process by which we replace the existing set of status symbols—those which emphasize the acquisition of material wealth—with an altogether different set of social standards in which special status, to the extent that it continues to exist at all, is based on social responsibility, service to the community, and warm, decent, and honorable relationships with one's fellow human beings.

¹⁷ Nor is it likely that a capitalist economy could function otherwise. A zero rate of growth means either that property owners must consume their profits, a difficult task to say the least, given the existing distribution of income—and one not likely to receive social sanction—or that profits be distributed directly in the form of gifts or loans or indirectly in the form of lower prices in which case no rationale remains for continuation of a private ownership system.

¹⁸ A. A. Berle (p. 102), a former State Department official and professor of corporation law at Columbia University, does not exaggerate when he writes:

But in terms of power, without regard to asset positions, not only do 500 corporations control two-thirds of the non-farm economy but within each of that 500 a still smaller group has the ultimate decision-making power. This is, I think, the highest concentration of economic power in recorded history. Since the United States carries on not quite half of the manufacturing production of the entire world today, these 500 groupings—each with its own little dominating pyramid within it—represent a concentration of power over economics which makes the medieval feudal system look like a Sunday school party. In sheer economic power this has gone far beyond anything we have yet seen.

Berle, of course, draws other, more conservative conclusions than I from this perception of social reality.

¹⁹ See J. O'Connor (1968) for a more elaborate statement of this argument and its implications.

that initially large firms achieved an even larger share in the later year.

To facilitate an analysis as to whether biases have crept into my procedures, I have classified those firms that exit from the ranks of the largest 100 by the ranks they previously had held and entrants by the ranks they achieve. For 1909-1919, we find that of those firms ranked 1-50 in 1909, 19 were ranked 51 or smaller in 1919. Of these, 5 were "exiters" while the other 14 managed to survive in the lower half of the 1919 listings. As many as 34 firms, ranked 51-100 in 1909, dropped out of the largest 100 by 1919. Thus, most of the non-surviving firms of 1909-1919 were of small or medium size. Taking into account only the "exiting" firms of 1909-1919, it would appear that the survivors only method understates the regression coefficient. This can be explained as follows: most of the 1909 firms—14 of 19—which fall out of the ranks of the largest 50 affect the regression since they remain among the survivors. None of the 36 smaller firms which drop into ranks 101-150 affect the regression. The regression coefficient is therefore *understated* since the relative declines of the large firms are counted while those of the smaller firms are not.

In contrast, smaller firms not ranked in 1909 but ranked among the largest 100 of 1919 often made gains relative to the larger firms. Fifteen of the 39 entrants of 1909-1919 achieved a position in the first 50 ranks. To the extent that these gains of the relatively small do not register in the 1909-1919 regression, the regression coefficient may be *overstated*. We therefore conclude that

the survivors only method unavoidably introduces a set of biases for the period 1909-1919 of undetermined direction and magnitude, but they appear to work at cross purposes and thereby tend to cancel each other out.

The situation as to entries and exits, as revealed in Table 1, changes little between the first decade and the second, but after 1929, the picture is considerably different. Rarely does an entrant rise into the ranks of the largest 50; rarely does a member of the largest 50 fall out of the ranks of the largest 100. As a result, a small downward bias probably exists; that is, the late coefficients are somewhat lower than they should be. The reason is somewhat complicated. When exiters are small, the gains of the large firms vis-à-vis the small exiters are not understated. In contrast, the gains made by the entrants are at the expense of the smaller firms—those ranked 50 or higher—and not at all at the expense of the largest corporations. For this reason, the regression coefficient is relatively unaffected. To the extent the above analysis is correct, it is possible that we may have slightly understated the increased ability of the large firms to maintain their shares over the years 1909-1964.

What we are trying to determine is whether the survivors only method which uses most of the 100 largest firms, but not all of them, cause *b* the regression coefficient to be either understated or overstated. This problem tends to diminish as more and more firms are used as observations. Since this is true, we can look at the path taken by the regression coefficients as we move from a study

TABLE 1—EXITING AND ENTERING FIRMS, CLASSIFIED BY RANKS

		Ranks										Total
		1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	
1909-1919	Exiters	0	1	0	2	2	4	7	9	8	6	39
	Entrants	4	1	3	4	3	6	4	3	5	6	39
1919-1929	Exiters	1	1	1	2	2	3	5	3	6	6	30
	Entrants	0	2	1	4	3	3	6	3	5	3	30
1929-1939	Exiters	0	1	0	0	3	0	3	2	2	4	15
	Entrants	0	0	0	0	0	1	2	1	5	6	15
1939-1948	Exiters	0	0	0	0	0	0	1	6	4	5	16
	Entrants	0	1	1	0	1	1	1	4	3	4	16
1948-1958	Exiters	0	0	0	0	0	0	1	3	6	5	15
	Entrants	0	0	0	2	2	2	0	6	1	2	15
1958-1964	Exiters	0	0	0	0	0	0	0	2	4	6	12
	Entrants	0	0	0	1	0	0	1	1	3	6	21

TABLE 2—REGRESSION COEFFICIENTS

	1909-1919	1919-1929	1929-1939	1939-1948	1948-1958	1958-1964
All survivors	0.73293	0.65403	0.96479	0.92046	1.11028	0.97590
First 75	0.75447	0.60347	0.97295	0.91434	1.08627	0.96163
First 50	0.76234	0.57467	0.96977	0.89558	1.13776	0.97254
First 25	0.79053	0.54448	0.92041	0.85448	1.13080	0.91898

TABLE 3—REGRESSION COEFFICIENTS, ADJUSTED FOR CHANGE IN INDUSTRIAL STRUCTURE

	1909-1919	1919-1929	1929-1939	1939-1948	1948-1958	1958-1964
All survivors	0.77094	0.84343	0.95668	0.92068	1.06927	0.96762
First 75	0.75591	0.79136	0.98028	0.90512	1.05739	0.95544
First 50	0.76675	0.75349	0.99998	0.90908	1.11953	0.88738
First 25	0.90691	0.73752	0.96544	0.90825	1.11928	0.92386

based on survivors from the ranks of the leading 25 firms to one based on survivors from the leading 100. In Tables 2 and 3 the various regression coefficients are displayed in a form suitable for a comparison of these differences.

The data indicates that both the adjusted as well as the unadjusted regression coefficients for all survivors may be somewhat overstated in 1909-1919 and understated in 1919-1929; for the decades 1939-1948 and 1958-1964, there appears to be a slight understatement, while for 1948-1958 a slight overstatement. No bias one way or the other appears in the remaining decade. This information does not conflict with our earlier judgments; if anything, it confirms them. We had suggested that the bias was indeterminate for the first two decades and our "trend of the coefficients" estimate of the bias shows that in later decades the bias, if any, tended towards understatement of the regression coefficient. The above data shows that in two of the three decades for which a trend exists, the coefficient is slightly understated. With so few observations, too much should not be made of these results. All things considered, the bias introduced by my procedure is probably negligible and of limited significance. Still, to the extent we can check on the direction of the bias inherent in the survivors only method, the major conclusion of my article, that the regression coefficients have exhibited a distinct upward trend indicating increased ability of the large corporations to maintain their shares, goes unchallenged. In fact, it is reinforced, given the lowered coefficient for 1909-1919 and the raised coefficient for 1958-1964.

REFERENCES

- P. Baran, *The Political Economy of Growth*, New York 1957.
- and P. Sweezy, *Monopoly Capital: An Essay on the American Economic and Social Order*, New York 1966.
- A. A. Berle, "Economic Power and the Free Society," in A. Hacker, ed., *The Corporation Take-Over*, New York 1964.
- D. Chapman and P. Harrington, "Land Lovers," *Look*, November 4, 1969, 33, 54-61.
- B. Commoner, *Science and Survival*, New York 1963.
- G. W. Domhoff, *Who Rules America?* Englewood Cliffs 1967.
- J. K. Galbraith, *The Affluent Society*, Boston 1958.
- , *The New Industrial State*, Boston 1967.
- A. D. H. Kaplan, *Big Enterprise in a Competitive System*, Washington 1954; rev. ed., Washington 1964.
- H. Marcuse, *An Essay on Liberation*, Boston 1969.
- , *One-Dimensional Man*, Boston 1964.
- D. Mermelstein, *Economics: Mainstream Readings and Radical Critiques*, New York 1970.
- , "Large Industrial Corporations and Asset Share Maintenance, 1909-1964," unpublished doctoral dissertation, Columbia Univ. 1967.

- , "Large Industrial Corporations and Asset Shares," *Amer. Econ. Rev.*, Sept. 1969, 59, 531-41.
- R. Miliband, "Professor Galbraith and American Capitalism," in R. Miliband and J. Saville, eds., *The Socialist Register 1968*, London 1968.
- J. O'Connor, "The Fiscal Crisis of the State," Part I, *Socialist Revolution*, Jan.-Feb. 1970, 1, 14-53; Part II, Mar.-Apr. 1970, 2, 39-94.
- , "The Situation at Present and What Is To Be Done?" *Mid-Peninsula Observer*, June 3-17, 1968, 9-12; reprinted under the title, "Some Contradictions of Advanced U.S. Capitalism," in D. Mermelstein, ed., *Economics: Mainstream Readings and Radical Critiques*, New York 1970.
- S. R. Reid, *Mergers, Managers, and the Economy*, New York 1968.
- K. W. Rothschild, "Price Theory and Oligopoly," *Econ. J.* Sept. 1947, 57, 299-320; reprinted in K. E. Boulding and G. J. Stigler, eds., *Readings in Price Theory*, Homewood 1952.
- U.S. Congress, Subcommittee on Antitrust and Monopoly of the Committee on the Judiciary, U.S. Senate, *Economic Concentration*, Hearings, Pt. 1, 88th Cong., 2nd sess., Washington 1964.

Welfare Aspects of a Regulatory Constraint: Note

By EYTAN SHESHINSKI*

Government agencies commonly employ the "fair rate of return" criterion in the regulation of monopolies: after the firm subtracts its operating expenses from gross revenues, the remaining revenue should be just sufficient to compensate the firm for its investment in plant and equipment, at a rate which is considered to be fair.

It has been argued by Harvey Averch and Leland Johnson, and now rigorously proved by Akira Takayama, that such constraint induces the firm, subject to regulatory control, to increase its investment and output and also to deviate from the optimal allocation of inputs, because the regulated firm does not equate marginal rates of factor substitution to the ratio of factor costs. Therefore, cost is not minimized at the output selected by the firm.¹

Since the fair rate of return criterion leads to a nonoptimal state in the sense of Pareto, a basic question is whether it improves the performance of the economy, from a welfare point of view, as compared with the unregulated monopoly situation (where output is too small). This is a "second best" problem in which we have to choose between two situations, each deviating in one way or another from optimality.

Here we show that from the point of view of efficiency, disregarding income distribution aspects, some regulation via the fair rate of return is always advantageous. We also derive the rule for the optimal degree of regulation, i.e. the regulation that maximizes social welfare.

I. The Takayama Model

Consider a monopoly employing two inputs, capital K and labor L to produce a

homogeneous output Y . Output is a function of inputs

$$(1) \quad Y = f(K, L)$$

Each factor has a positive and decreasing marginal product. The marginal rate of substitution between inputs is decreasing, i.e., isoquants are concave.

The price of the product P is negatively related to the level of output by the inverse demand function

$$(2) \quad P = P(Y), \quad P'(Y) < 0$$

The costs of the inputs, r for capital and w for labor, are fixed for the firm. Total costs, C , are therefore $C = rK + wL$. The profit of the firm Π , is defined as

$$(3) \quad \Pi = PY - C = PY - rK - wL$$

Since depreciation is assumed away, the operating expenses of the firm are only labor costs. The fair rate of return criterion imposes the following constraint on the firm: Denote by s the fair rate of return determined by the regulatory agency. The firm's net revenues (gross revenues minus operating expenses) per unit of capital should not exceed s

$$\frac{P \cdot Y - wL}{K} \leq s$$

or

$$(4) \quad P \cdot Y - sK - wL \leq 0$$

The firm attempts to maximize (3) subject to (4). Under suitable assumptions about the profit function, and provided the solution is interior, the first-order necessary conditions for a maximum are²

$$(5) \quad (P + P'f)f_1 - r - \lambda[(P + P'f)f_1 - s] = 0$$

* Hebrew University, Jerusalem.

¹ This deviation from the optimal allocation of inputs persists even under a more flexible scheme for a "graduated fair return" criterion, as suggested by Alvin Klevorick, although the degree of inefficiency is clearly reduced.

² A rigorous analysis is given by Takayama.

$$(6) \quad (P + P'f)f_2 - w - \lambda[(P + P'f)f_2 - w] = 0$$

λ being the Lagrange multiplier.

When s exceeds the maximum rate of return earned by the unregulated monopoly, \bar{s} , the constraint is not binding, λ is zero, and we have from (5) and (6) the standard profit maximizing condition of a monopolist. At the other extreme, when s is less than r , profits are negative and the firm will prefer to shut down. In the border case $s=r$ (and $\lambda=1$) the firm is indifferent as between shutting down and operating. Thus, the realm of interest is that in which $0 < \lambda < 1$, where the constraint is effective ($\lambda > 0$) and the fair rate of return exceeds the market cost of capital ($\lambda < 1$).

At the constrained maximum point, production is inefficient, since from (5) and (6)

$$(7) \quad \frac{f_1}{f_2} = \frac{r}{w} - \frac{\lambda}{1-\lambda} \left(\frac{s-r}{w} \right) < \frac{r}{w}$$

the marginal rate of substitution between inputs is lower than the ratio of input prices. Each output is produced with more capital and less labor as compared to the unregulated optimum. This inefficiency derives from the fact that the net return of the firm on every unit of capital is $s-r$, and this creates an incentive to increase capital intensity.³

In the region in which $0 < \lambda < 1$, we have from (6)

$$(8) \quad (P + P'f)f_2 - w = 0$$

which is the standard profit maximizing rule of a monopoly for labor input. Equation (8), together with the constraint (4)

$$(9) \quad P \cdot Y - sK - wL = 0$$

determine the constrained inputs (K , L) for any given level of s . We now wish to analyze the response of the firm to different levels of regulation.

Differentiating the profit constraint (9) with respect to s ,

³ In the constrained region $\Pi = P \cdot Y - rK - wL = P \cdot Y - sK - wL + (s-r)K = (s-r)K$

$$(10) \quad (P + P'f) \left(f_1 \frac{dK}{ds} + f_2 \frac{dL}{ds} \right) - K - s \frac{dK}{ds} - w \frac{dL}{ds} = 0$$

From (8) this reduces to

$$(11) \quad [(P + P'f)f_1 - s] \cdot \frac{dK}{ds} = K$$

Since $s > r$ and $0 < \lambda < 1$, we have from (5) that $(P + P'f)f_1 - s < 0$, so (11) yields that $dK/ds < 0$. The response of L to changes in s is found by differentiation of (8). Denoting the marginal revenue function by $h = P + P'f$, we have

$$(12) \quad [h'f_1f_2 + hf_{21}] \cdot \frac{dK}{ds} + [h'f_2^2 + hf_{22}] \cdot \frac{dL}{ds} = 0$$

It is assumed that marginal revenue is decreasing, $h' < 0$, so the sign of the second-term in brackets is negative. Since $dK/ds < 0$, dL/ds is negative provided the first-term in brackets is positive:

$$(13) \quad h'f_1f_2 + hf_{21} > 0$$

Let us make this assumption, which is equivalent to assuming that capital and labor are complement inputs.⁴

Thus, as regulation tightens, i.e., as s decreases, both inputs and output increase.

$$-\frac{dK}{ds} > 0, \quad -\frac{dL}{ds} > 0, \quad -\frac{dY}{ds} > 0$$

II. Welfare Implications

Suppose the economy consists of identical individuals, so there is no income distribution problem. The utility, or social welfare function, is

$$(14) \quad U = U(Y, K, L)$$

The marginal utility of Y is positive,

⁴ It can be shown that $\partial K/\partial w < 0$ and $\partial L/\partial r < 0$. Inputs are normal and complements provided (in addition to the other assumptions), $\partial^2 \pi / \partial K \partial L = h'f_1f_2 + hf_{21} > 0$.

$U_1 > 0$, while the marginal utilities of capital and labor are negative, $U_2 < 0$ and $U_3 < 0$, reflecting the loss of forgone consumption and leisure, respectively.

The change in utility as a result of a change in s is

$$(15) \quad \frac{dU}{ds} = U_1 \left(f_1 \frac{dK}{ds} + f_2 \frac{dL}{ds} \right) + U_2 \frac{dK}{ds} + U_3 \frac{dL}{ds}$$

The first-order conditions of the consumer who maximizes utility are

$$(16) \quad \frac{U_1}{P} = - \frac{U_2}{r} = - \frac{U_3}{w}$$

Substituting in (15) we get

$$(17) \quad \frac{dU}{ds} = \frac{U_1}{P} \cdot \left[(Pf_1 - r) \frac{dK}{ds} + (Pf_2 - w) \frac{dL}{ds} \right]$$

Let us determine the sign of dU/ds at the unconstrained point: $s = \bar{s} (\lambda = 0)$. At this point, from (5) and (6)

$$Pf_1 - r = -P'ff_1$$

and

$$Pf_2 - w = -P'ff_2$$

Substituting in (17) we have

$$(18) \quad \frac{dU}{ds} = - \frac{U_1}{P} P'f \cdot \left(f_1 \frac{dK}{ds} + f_2 \frac{dL}{ds} \right) < 0$$

Decreasing s from the ineffective level \bar{s} , always raises utility. Therefore, *some regulation via the fair rate of return criterion is always worthwhile.*

III. Optimal Degree of Regulation

Since regulation can always improve welfare, it is interesting to find the level of s that maximizes utility. In the constraint region ($r < s < \bar{s}$), the necessary condition for

a maximum of U is to set (18) equal to zero. This condition can be rewritten

$$(19) \quad \frac{dU}{ds} = \frac{U_1}{P} \cdot \left[P \frac{dY}{ds} - \frac{dC}{ds} \right] = 0$$

or

$$(20) \quad P \frac{dY}{ds} = \frac{dC}{ds}$$

where C is total costs. Condition (20) has an obvious interpretation: as s decreases, output and costs increase. As we have seen, the increase in costs stems both from the increase in output and from the inefficient combination of inputs.

It is desired to reduce s so long as the resulting value of the increased output exceeds the corresponding increase in costs, and vice versa. At the optimum point, the change in the value of output exactly equals the change in costs. Put in a different way, (20) can be rewritten as

$$(21) \quad P = \frac{dC}{dY}$$

which resembles the standard optimality rule of price equal to marginal costs, only that dC/dY is not the curve derived from minimum cost allocation of inputs.

In Figure 1, MC is the marginal cost curve

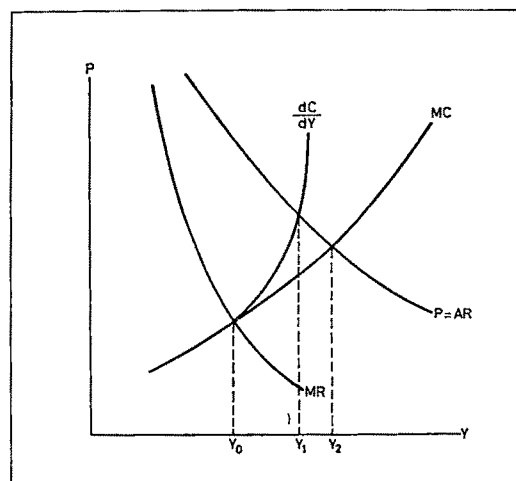


FIGURE 1

of the *unconstrained* monopoly, P the demand curve and MR the marginal revenue curve. The unconstrained optimal output is Y_0 . Now, as s decreases from s , output and costs increase. The change in costs for a unit change in output is drawn by the curve dC/dY which is higher throughout than the MC curve (which represents minimum combination of inputs). The optimal s is set at the level corresponding to the point where dC/dY intersects the demand curve, at output Y_1 .⁵ The optimal output is always higher

than that produced by the unconstrained monopoly (Y_0), and lower than the Pareto optimum output (Y_2).

REFERENCES

- H. Averch and L. L. Johnson, "Behavior of the Firm Under Regulatory Constraint," *Amer. Econ. Rev.*, Dec. 1962, 52, 1052-69.
 A. K. Klevorick, "The Graduated Fair Return: A Regulatory Proposal," *Amer. Econ. Rev.*, June 1966, 56, 477-84.
 A. Takayama, "Behavior of the Firm Under Regulatory Constraint," *Amer. Econ. Rev.*, June 1969, 59, 255-60.

⁵ A comment on the "graduated fair return" criterion, suggested by Klevorick, might be in place here. He proposes that the fair return be a function of the amount of capital that the firm employs, rather than a fixed number. If one can choose the function optimally, a graduated maximal return can clearly not be worse than a fixed one, since the latter is a special case of the former. While this much is obvious, the index chosen by Klevorick, E , which indicates the deviation of the mar-

ginal rate of substitution between factors and the ratio of their prices, has no direct welfare implications. In fact, the absolute deviation of E from unity (Pareto optimality) is monotonically increasing as s decreases, but we already know that not all reductions in s necessarily improve welfare.

Pitfalls in Financial Model Building: Some Extensions

By MARK L. LADENSON*

In a recent contribution to this *Review*, William Brainard and James Tobin presented a general equilibrium model of the financial sector of the economy and a general disequilibrium model of the dynamic process by which the endogenous variables in the model adjust from one set of equilibrium values to another in response to a change in one or more exogenous variables. The variables in the models are linked by behavioral equations and by identities. The identities imply a number of restrictions on the coefficients of the variables in the behavioral equations and in general, imply that the coefficients of a particular variable summed across the behavioral equations add up to zero. This was one of the authors' basic points but the argument was presented without any formal derivations of the restrictions. Furthermore, the authors did not estimate the coefficients of the equations of their model but assumed certain values for the parameters and performed simulation experiments. In Section I of the present essay we give a formal derivation of the key restrictions of Brainard and Tobin's model, and in Section II we discuss some problems associated with estimating the parameters of the model.

I

Brainard and Tobin's static equilibrium model includes a sector containing financial assets and debts of the public and a sector involving bank asset holdings. In deriving the restrictions emphasized by the authors, however, it is unnecessary to deal with the banking sector. Brainard and Tobin assume that the public holds its net wealth in the

form of five different types of financial assets and liabilities: demand deposits, time deposits, treasury securities, loans from banks, and equities. The demand for each of these is a linear function, homogeneous in net worth, of interest rates and national income:

$$(1) \quad y_t^* = (B_1 X_1 + B X_t) W_t$$

where X_1 is a variable which always takes on a value of unity, X_t is a 6x1 vector of interest rates and national income, B_1 is the 5x1 vector of coefficients of X_1 (constant terms), B is the 5x6 matrix of structural coefficients, W_t is net worth, a scalar, and y_t^* is the 5x1 vector of the values of the financial assets and liabilities given by current values of X_t and W_t .

The static equilibrium model of public behavior contains the system of five equations (1), and implicitly it also contains an *ex post* and an *ex ante* identity. Define the vector:

$$r' \equiv [1, 1, 1, -1, 1]$$

The *ex post* identity is

$$(2) \quad r' y_t = W_t$$

where y_t is a 5x1 vector of actual values of the financial assets and liabilities of the public sector.¹ Equation (2) is simply a balance sheet identity; the sum of assets minus liabilities equals net worth.

The *ex ante* identity

$$(3) \quad r' y_t^* = W_t$$

is a sort of "rational desires" hypothesis. It constrains the *desired* or equilibrium values of the financial assets and liabilities to obey the balance sheet identity and is analogous to the budget constraint in the familiar static theory of consumer behavior.

¹ The fourth element in the vector y_t , loans from banks, is the only liability of the public; hence the negative value of the fourth element of r' .

* Assistant professor of economics, Michigan State University. The chapter of my doctoral dissertation on which this paper is based was written with major assistance from Professor Walter D. Fisher. Others whose help and influence are reflected in the paper are Professors Frank Brechling, Patric Hendershott, Jan Kmenta, James Ramsey, and Arthur Treadway. While my deep thanks go to them, I remain solely responsible for errors.

$$(4) \quad \begin{array}{l} y_{1t} - y_{1,t-1} = \alpha_{11}(y_{1t}^* - y_{1,t-1}) + \dots + \alpha_{15}(y_{5t}^* - y_{5,t-1}) + \gamma_{1\Delta W}\Delta W_t \\ \vdots \\ y_{5t} - y_{5,t-1} = \alpha_{51}(y_{5t}^* - y_{5,t-1}) + \dots + \alpha_{55}(y_{5t}^* - y_{5,t-1}) + \gamma_{5\Delta W}\Delta W_t \end{array}$$

Brainard and Tobin go on to specify how this system behaves out of equilibrium with the set of adjustment equations (4), where the α_{ii} are own-adjustment coefficients, the α_{ij} ($i \neq j$) are cross-adjustment coefficients, the $\gamma_{i\Delta W}$ are structural coefficients, and the subscript on y indexes a particular financial asset (or liability). The system (4) is a general disequilibrium system, a generalization of the familiar stock-adjustment model. The change in each financial asset depends not only on the gap between its own desired and actual values but on all such gaps and also on the change in wealth in the period. Equation (4) may be expressed more compactly as

$$(5) \quad y_t - y_{t-1} = A(y_t^* - y_{t-1}) + \gamma_{\Delta W}\Delta W_t$$

where A is the 5×5 matrix of adjustment coefficients, α_{ij} , and $\gamma_{\Delta W}$ is the 5×1 vector of structural coefficients, the i th element of which is $\gamma_{i\Delta W}$. Adding the vector y_{t-1} to both sides gives

$$(6) \quad y_t = Ay_t^* + Zy_{t-1} + \gamma_{\Delta W}\Delta W_t$$

where $Z = I_5 - A$, with I_5 denoting an identity matrix of order 5. Equations (1), (2), (3), and (6) will be called Model I. They constitute a system of twelve equations in ten endogenous variables (the elements of the vectors y^* and y). There are more equations than unknowns and the question arises as to whether this system is consistent. A sufficient condition for the existence of a solution is given by the following well-known theorem: "A system $CX + b = 0$ of m linear equations in n unknowns is consistent if, and only if, the coefficient matrix C and the augmented matrix $[Cb]$ have the same rank" (Franz Hohn p. 140.) Treating the elements of the vector y as the first five unknowns and the elements of the vector y^* as the next five and writing the twelve equations in the fol-

lowing order—(6), (1), (3), (2)—the 12×10 coefficient matrix, C , of the system is

$$(7) \quad C = \begin{bmatrix} I_5 & -A \\ 0 & I_5 \\ r' & 0 \\ 0 & r' \end{bmatrix}$$

The 12×11 augmented matrix $[Cb]$ of the system then is

$$(8) \quad [Cb] = \begin{bmatrix} I_5 & -A & -(Zy_{t-1} + \gamma_{\Delta W}\Delta W_t) \\ 0 & I_5 & -(B_1X_1 + BX_t)W_t \\ r' & 0 & -W_t \\ 0 & r' & -W_t \end{bmatrix}$$

According to the theorem, sufficient conditions for these two matrices to have the same rank are also sufficient conditions for the system of equations, Model I, to be consistent. We therefore seek sufficient conditions for the rank of the matrix (8) to have the same rank as the matrix (7). We may form rank-equivalent matrices by elementary row operations. This will be facilitated by partitioning the matrix (8) as follows:

$$(8') \quad [Cb] = \begin{bmatrix} M_{1-5} \\ M_{6-10} \\ M_{11} \\ M_{12} \end{bmatrix}$$

where M_{1-5} represents the first five rows of (8), M_{6-10} represents the next five, M_{11} the eleventh, and M_{12} the twelfth row. We now postmultiply r' by M_{1-5} and subtract the result from M_{11} to get

$$(9) \quad \begin{bmatrix} I_5 & -A & -(Zy_{t-1} + \gamma_{\Delta W}\Delta W_t) \\ 0 & I_5 & -(B_1X_1 + BX_t)W_t \\ 0 & r'A & -W_t + r'(Zy_{t-1} + \gamma_{\Delta W}\Delta W_t) \\ 0 & r' & -W_t \end{bmatrix}$$

$$(10) \quad \begin{bmatrix} I_5 & -A & -(Zy_{t-1} + \gamma_{\Delta W} \Delta W_t) \\ 0 & I_5 & -(B_1 X_1 + B X_t) W_t \\ 0 & 0 & -W_t + r'(Zy_{t-1} + \gamma_{\Delta W} \Delta W_t) + r' A (B_1 X_1 + B X_t) W_t \\ 0 & 0 & -W_t + r'(B_1 X_1 + B X_t) W_t \end{bmatrix}$$

Partitioning (9) as follows:

$$(9') \quad \begin{bmatrix} N_{1-5} \\ N_{6-10} \\ N_{11} \\ N_{12} \end{bmatrix}$$

where the N_i represent row(s) of (9) and the subscripts are to be interpreted as in (8'), we postmultiply $r'A$ by N_{6-10} and subtract the result from N_{11} , and postmultiply r' by N_{6-10} and subtract the result from N_{12} , to get the matrix that we have denoted as (10). The matrix (10) has the same rank as the matrix (8) and will have the same rank as the matrix (7) if the elements in the eleventh column of the last two rows of this matrix are equal to zero; that is, if

$$(11) \quad \begin{aligned} W_t &= r'(Zy_{t-1} + \gamma_{\Delta W} \Delta W_t) \\ &+ r' A (B_1 X_1 + B X_t) W_t, \end{aligned}$$

and

$$(12) \quad W_t = r'(B_1 X_1 + B X_t) W_t$$

If (11) and (12) are satisfied, our system of twelve equations in ten unknowns will have at least one solution. Considering first the sufficient condition for (12) to be satisfied, we recall that X_1 is a scalar, identically equal to one. Clearly, therefore, (12) will be satisfied if

$$(12a) \quad \begin{cases} r' B_1 = 1 \\ r' B = 0 \end{cases}$$

where 0 is a row vector of six zeros. Conditions (12a) embody the restrictions discussed by Brainard and Tobin (pp. 103, 107) on the structural coefficients of the static equilibrium model. These conditions require constant terms summed over equations (1) to

equal unity and coefficients of any interest rate or income so summed to add to zero. If a change in an exogenous variable (interest rate, income) induces an increase in the desired amount of one asset, it induces corresponding decreases (increases) in the desired amounts of some other assets (liabilities). This is the basis of the authors' plea that all interest rates and income should be entered as exogenous variables in the equation of each asset and liability.

Turning next to sufficient conditions for (11) to be satisfied, we first substitute for Z to get:

$$\begin{aligned} W_t &= r' A (B_1 X_1 + B X_t) W_t \\ &+ r' (I_5 - A) y_{t-1} + r' \gamma_{\Delta W} \Delta W_t \end{aligned}$$

Subtracting $r'y_{t-1}$ from both sides gives:

$$\begin{aligned} W_t - r'y_{t-1} &= r' A [(B_1 X_1 + B X_t) W_t \\ &- y_{t-1}] + r' \gamma_{\Delta W} \Delta W_t \end{aligned}$$

In view of (2) we may write:

$$(11') \quad \begin{aligned} \Delta W_t &= r' A [(B_1 X_1 + B X_t) W_t \\ &- y_{t-1}] + r' \gamma_{\Delta W} \Delta W_t \end{aligned}$$

Clearly one sufficient set of conditions for (11') to be satisfied is:

$$(11a) \quad \begin{cases} r' A = 0 \\ r' \gamma_{\Delta W} = 1 \end{cases}$$

where 0 is a row vector of five zeros. Conditions (11a) embody the restrictions discussed by Brainard and Tobin (pp. 106, 108) on the adjustment coefficients of the dynamic system (6). In strict analogy to the case of the structural coefficients, if a gap between desired and actual amounts of a particular asset induces an increase in the actual value of that asset, it induces corresponding de-

creases (increases) in the actual values of other assets (liabilities). Conditions (11a) require the adjustment coefficients summed over a particular gap to add to zero and require the coefficients of the ΔW term to add to unity (as one would surely expect they must since the proportions in which a change in net worth is distributed over assets and liabilities add to 100 percent).

Brainard and Tobin stated that conditions (11a) are necessary and sufficient for satisfaction of (11'). It is clear, however, that if (12a) is satisfied, they are not necessary. An alternative set of conditions exists:

$$(11b) \quad \begin{cases} r'A = r' \\ r'\gamma_{\Delta W} = 0 \end{cases}$$

as is easily verified by first substituting (11b) into (11'), noting the restrictions on B_1 and B given in (12a), and noting that by equation (2) $r'y_{t-1} = W_{t-1}$. Conditions (11b) require adjustment coefficients summed over an asset (liability) gap to equal unity (minus one) and require the coefficients of the ΔW term to sum to zero rather than unity. In this case the requirement that the proportions in which a change in wealth is distributed over assets and liabilities add to 100 percent is taken care of by the use of the restriction on the vector B_1 given in conditions (12a).

II

Brainard and Tobin assumed particular values for the elements of the A and B matrices and the $\gamma_{\Delta W}$ vector. However, one might well wish to estimate the system (6) empirically. Several problems arise in this connection. They will be discussed in turn.

We begin by noting that the equation system (6) is deterministic. Treating the scalar W_t and the elements of the vector X_t as fixed in repeated samples we may introduce stochastic elements into it:

$$(6v) \quad y_t = Ay_t^* + Zy_{t-1} + \gamma_{\Delta W}\Delta W_t + v_t,$$

where v_t is a 5×1 vector of disturbance terms and we assume that its elements have zero expectation and serial independence and are homoskedastic.

Our complete model now consists of the equations (1), (2), (3), and (6v), which is again a system of twelve equations in ten endogenous variables. We will call this system Model II. It differs from Model I in that the five equations (6) have been replaced by the five equations (6v). Sufficient conditions for consistency of Model II are derived in the same manner used in deriving such conditions for consistency of Model I. They require satisfaction of the following equations:

$$(11v) \quad -W_t + r'(Zy_{t-1} + \gamma_{\Delta W}\Delta W_t + v_t) + r'A(B_1X_1 + BX_t)W_t = 0$$

$$(12) \quad -W_t + r'(B_1X_1 + BX_t)W_t = 0$$

We have already seen that equation (12) is satisfied by the conditions (12a). To find sufficient conditions for (11v) to be satisfied we proceed as with equation (11). Substituting for Z , subtracting $r'y_{t-1}$ from both sides, and writing ΔW_t for $W_t - r'y_{t-1}$ we get:

$$(11v') \quad \begin{aligned} \Delta W_t &= r'A[(B_1X_1 + BX_t)W_t \\ &\quad - y_{t-1}] + r'(\gamma_{\Delta W}\Delta W_t + v_t) \end{aligned}$$

Clearly one sufficient set of conditions for (11v') to be satisfied is:

$$(11v, a) \quad \begin{cases} r'A = 0 \\ r'\gamma_{\Delta W} = 1 \\ r'v_t = 0 \end{cases}$$

where 0 is a row vector of five zeros. However an alternative set of sufficient conditions exists:

$$(11v, b) \quad \begin{cases} r'A = r' \\ r'\gamma_{\Delta W} = 0 \\ r'v_t = 0 \end{cases}$$

as is easily verified by first substituting (11v, b) into (11v'), noting the restrictions on B_1 and B given in (12a), and noting that by equation (2) $r'y_{t-1} = W_{t-1}$. Conditions (11v, a) and (11v, b) differ from conditions (11a) and (11b), respectively, in that they further require $r'v_t = 0$. That is, they require the elements of the disturbance vector to sum to zero. The other features of conditions

(11a) and (11b) are preserved in Model II. We now proceed to an interpretation of the difference between conditions (11v, a) and conditions (11v, b).

Substituting (1) into (6v) gives the following system:

$$\begin{aligned}
 y_t &= A(B_1 X_1 + B X_t) W_t + Z y_{t-1} \\
 &\quad + \gamma_{\Delta W} \Delta W_t + v_t \\
 (13) \quad &= [AB_1 \ AB \ Z \ \gamma_{\Delta W}] \begin{bmatrix} W_t \\ X_t W_t \\ y_{t-1} \\ \Delta W_t \end{bmatrix} + v_t
 \end{aligned}$$

The system of reduced equations for estimating the parameter matrices A and B and vectors B_1 and $\gamma_{\Delta W}$ is given by

$$(13R) \quad y_t = [\gamma_1 \ \Gamma_{2,7} \ \Gamma_{8,12} \ \gamma_{\Delta W}] \begin{bmatrix} W_t \\ X_t W_t \\ y_{t-1} \\ \Delta W_t \end{bmatrix} + v_t$$

where γ_1 is the 5×1 vector of reduced form coefficients of W_t , $\Gamma_{2,7}$ is the 5×6 matrix of reduced form coefficients of the exogenous variables, $\Gamma_{8,12}$ is the 5×5 matrix of reduced form coefficients of the lagged endogenous variables, and $\gamma_{\Delta W}$ is the 5×1 vector of coefficients of the ΔW term. But unique estimates of γ_1 , $\Gamma_{8,12}$ and $\gamma_{\Delta W}$ cannot be obtained since an exact linear relation holds between W_t , ΔW_t and the five elements of y_{t-1} :

$$(2') \quad W_t = \Delta W_t + r' y_{t-1}$$

The problem can be made explicit. We may substitute (2') into (13R) to eliminate W_t from the latter.² Upon collecting terms we get

$$\begin{aligned}
 y_t &= [\Gamma_{2,7}, \Gamma_{8,12} + \gamma_1 r', \gamma_1 + \gamma_{\Delta W}] \\
 (13Ra) \quad &\cdot \begin{bmatrix} X_t W_t \\ y_{t-1} \\ \Delta W_t \end{bmatrix} + v_t
 \end{aligned}$$

² Of course we may also substitute (2') into the term $X_t W_t$ but no purpose is served by doing so, since that term has nothing to do with the problem under discussion.

It is quite clear that the individual components of the two sums $\Gamma_{8,12} + \gamma_1 r'$ and $\gamma_1 + \gamma_{\Delta W}$ are not identifiable. However, if we are willing to adopt the expedient of setting all components of the vector γ_1 equal to zero, the elements of $\Gamma_{8,12}$ and $\gamma_{\Delta W}$ are identifiable.

Instead of using (2') to eliminate W from (13R) we might just as easily have used it to eliminate ΔW . In that case we would get

$$\begin{aligned}
 y_t &= [\Gamma_{2,7}, \Gamma_{8,12} - \gamma_{\Delta W}, \gamma_1 + \gamma_{\Delta W}] \\
 (13Rb) \quad &\cdot \begin{bmatrix} X_t W_t \\ y_{t-1} \\ W_t \end{bmatrix} + v_t
 \end{aligned}$$

We cannot separately identify the components of $\Gamma_{8,12} - \gamma_{\Delta W}$ and of $\gamma_1 + \gamma_{\Delta W}$ but if we set all the elements of the vector $\gamma_{\Delta W}$ equal to zero, the elements of $\Gamma_{8,12}$ and of γ_1 are identifiable.³

To assume that the elements of γ_1 are all equal to zero and all the elements of X_t are not constant over time implies that W_t does not enter equation system (1) as an isolated variable in one of the linear terms. This in turn implies there is no functional dependence of the elements of the vector of equilibrium values of assets and liabilities, y_t^* , on net worth, apart from a scale factor. The first difference of net worth continues to enter the adjustment equations (6v), however, and does so without a coefficient of lagged adjustment. Therefore, to omit the variable W_t from system (1), while retaining ΔW_t in (6v) is to assume that the endogenous variables adjust instantaneously to a change in net worth.

On the other hand, to assume that the elements of $\gamma_{\Delta W}$ are all equal to zero is to omit the variable ΔW_t from equation system (6). If this is done it would appear that changes in assets and liabilities do not depend on net worth. However, since W_t continues to enter equations (1) which determine y_t^* , and since changes in assets and liabilities do depend on y_t^* , they also de-

³ Of course we can also achieve identifiability in a large number of other ways. Among other possibilities we might set any column of $\Gamma_{8,12}$ equal to zero.

pend on net worth. Furthermore, since these changes represent only partial adjustment to changes in y^* , they represent only partial (or lagged) adjustment to changes in net worth.

It follows that if W_t is dropped from the reduced equation system (13R) and ΔW_t enters these equations, instantaneous adjustment to a change in net worth is assumed and the conditions (11v,a) apply. If ΔW_t is omitted, W_t enters these equations, lagged adjustment to a change in net worth is assumed, all elements of the vector $\gamma_{\Delta W}$ are zero, and the conditions (11v, b) apply.

We must next show how we derive estimates of the elements of A , B , and B_1 from estimates of the parameters of the system (13R). Comparing (13) with (13R) it is seen that estimates of A are easily obtained. Since $Z = I_5 - A$, $A = I_5 - \Gamma_{8,12}$, and

$$(14) \quad \hat{A} = I_5 - \hat{\Gamma}_{8,12}$$

where hats indicate estimated values. Careful comparison of the two systems also shows that

$$\begin{aligned} \Gamma_{2,7} &= AB \\ \gamma_1 &= AB_1 \end{aligned}$$

Consider the case in which W is dropped from the regression equations. The elements of γ_1 are all equal to zero and estimates of the elements of $\gamma_{\Delta W}$ are given directly by the regression coefficients of ΔW . Estimates of A are obtained using (14). Denote this estimated matrix of adjustment coefficients as \hat{A}^a , and denote the matrix of estimates of $\Gamma_{2,7}$ as $\hat{\Gamma}_{2,7}^a$. Then

$$(15a) \quad \hat{\Gamma}_{2,7}^a = \hat{A}^a \hat{B}^a$$

where \hat{B}^a is an estimate of the matrix B and the superscript a indicates that W has been omitted from the relations. It would seem that the solution for \hat{B}^a is straightforward until one recalls the conditions (11a). They imply that the matrix A is singular. Assuming the estimate of A , \hat{A}^a , satisfies conditions (11a) it also is singular and no unique solution would seem to exist for \hat{B}^a . However we may add conditions (12a) to the system (15a) as follows:

$$(15a') \quad \begin{bmatrix} \hat{\Gamma}_{2,7}^a \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{A}^a \\ r' \end{bmatrix} \hat{B}^a$$

where 0 is a 1×6 row vector of zeros. Deleting all but one column from the left-hand matrix and all but the same column from \hat{B}^a , we get:

$$(15a'') \quad \begin{bmatrix} \hat{\Gamma}_j^a \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{A}^a \\ r' \end{bmatrix} [\hat{B}_j^a]$$

where the left-hand matrix denotes the j th column of the left-hand matrix of (15a') and \hat{B}_j^a denotes the j th column of \hat{B}^a . Then (15a'') is a system of six equations in five unknowns. We can find the requirement for this system to be consistent in the same way as we proceeded in the case of the system (8). The requirement is that $r' \hat{\Gamma}_j^a = 0$. If our estimate of Γ_j , $\hat{\Gamma}_j^a$, satisfies this condition, then by removing a dependent row of elements with row index i from \hat{A}^a and the element with the same row index from $\hat{\Gamma}_j^a$, the system is transformed into one of five independent equations which may be solved for the five unknowns \hat{B}_j^a .⁴ The process may be repeated five times until a solution for the entire matrix \hat{B}^a has been obtained.

If ΔW has been omitted from the equations, the elements of $\gamma_{\Delta W}$ are all equal to zero and estimates of the elements of A , \hat{A}^b say, are obtained using (14). If $\hat{\gamma}_1$ is the vector of estimates of the elements of γ_1 , and $\hat{\Gamma}_{2,7}^b$ is the matrix of estimates of the elements of $\Gamma_{2,7}$, then we wish to solve the system

$$(15b) \quad \hat{\gamma}_1 \hat{\Gamma}_{2,7}^b = \hat{A}^b [\hat{B}_1 \hat{B}^b]$$

for $[\hat{B}_1 \hat{B}^b]$, an estimate of $[B_1 B]$. Since conditions (11b) do not imply the singularity of A , the assumption that \hat{A}^b satisfies these conditions does not present the same difficulty as in the preceding paragraph, and the solution process is straightforward. However, the solution should be consistent with condi-

⁴ If the rank of \hat{A}^a is four (or more generally, if the rank is one less than the order of \hat{A}^a), consistency of (15a'') guarantees uniqueness of the solution, \hat{B}_j^a , regardless of the choice of the dependent row index i . I am indebted to M. J. Ringo for spotting an error on this point in an earlier draft.

tions (12a). To investigate the requirements for such consistency we form

$$(15b') \quad \begin{bmatrix} \hat{\gamma}_1 & \hat{\Gamma}_{2,7}^b \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \hat{A}^b \\ r' \end{bmatrix} [\hat{B}_1 \hat{B}^b]$$

and proceed to analyze it in the manner that system (8) was analyzed. The requirements for consistency turn out to be

$$\begin{aligned} r' \hat{\Gamma}_{2,7}^b &= 0 \\ r' \hat{\gamma}_1 &= 1 \end{aligned}$$

Hence our estimates of $\Gamma_{2,7}$ and of γ_1 must satisfy these conditions, if the solution for $[\hat{B}_1 \hat{B}^b]$ is to satisfy both conditions (11b) and (12a).

Finally, the question arises as to whether ordinary least squares estimates of $\Gamma_{8,12}$ yield estimates of A (by the transformation (14)) which satisfy (11a) or (11b) as may be appropriate, or whether some constrained estimation technique is necessary. The answer is that so long as all lagged endogenous variables appear in all equations of (13R), the ordinary least squares estimator of $\Gamma_{8,12}$ can be transformed into an estimate of A which satisfies the restrictions. We first prove the proposition for the case in which W is dropped. Our data are constrained by (2') which may be rewritten

$$r' y_t = r' y_{t-1} + \Delta W_t$$

or

$$r' y_t = \begin{bmatrix} 0 & r' & 1 \end{bmatrix} \begin{bmatrix} X_t W_t \\ y_{t-1} \\ \Delta W_t \end{bmatrix}$$

where 0 is a 1×6 vector of zeros, or

$$(16a) \quad r' y_t = q' f_a$$

$$\text{where } q' = \begin{bmatrix} 0 & r' & 1 \end{bmatrix} \text{ and } f_a = \begin{bmatrix} X_t W_t \\ y_{t-1} \\ \Delta W_t \end{bmatrix}$$

Transposing (16a) and writing it repeatedly for all T observations yields

$$(16a') \quad Yr = F_a q$$

where Y and F_a are the $T \times 5$ and $T \times 12$ data matrices on the variables. Assuming the elements of the vector γ_1 are all equal to zero, the ordinary least squares estimates of the parameters of equations (13Ra) are given by

$$(17a) \quad \hat{\Gamma}^a = (F_a' F_a)^{-1} F_a' Y$$

where Γ^a is the parameter matrix $[\Gamma_{2,7}, \Gamma_{8,12}, \gamma_{\Delta W}]$. Postmultiplying (17a) by r , and substituting Yr from (16a') yields

$$\hat{\Gamma}^a r = (F_a' F_a)^{-1} F_a' F_a q$$

or

$$\hat{\Gamma}^a r = q$$

Transposing, one gets

$$r' \hat{\Gamma}^a = q'$$

By comparing system (13) with (13Ra) one sees that $\hat{\Gamma}^a$ is an estimate of the parameter matrix $[AB \ Z \ \gamma_{\Delta W}]$. Making this substitution, and also substituting the definition of q' , one gets

$$(18a') \quad r' \hat{Z} = r'$$

$$(18a'') \quad r' A \hat{B} = 0$$

$$(18a''') \quad r' \gamma_{\Delta W} = 1$$

Since $\hat{Z} = I_5 - \hat{A}^a$, substituting into (18a') gives

$$r' (I_5 - A^a) = r' I_5$$

or $r' \hat{A}^a = 0$. Since we also have (18a'''), we have shown that our estimates satisfy conditions (11a).

Next consider the case in which ΔW is omitted. We rewrite (2') as

$$r' y_t = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ W_t \\ X_t W_t \end{bmatrix}$$

where the 0's represent 1×5 and 1×6 vectors of zeros, or

$$(16b) \quad r' y_t = p' f_b$$

where $p' = [0 \ 1 \ 0]$ and $f_b = \begin{bmatrix} y_{t-1} \\ W_t \\ X_t W_t \end{bmatrix}$

Transposing (16b) and writing it repeatedly for all T observations yields

$$(16b') \quad Yr = F_b q$$

where Y and F_b are the $T \times 5$ and $T \times 12$ data matrices on the variables. Assuming the elements of the vector $\gamma_{\Delta W}$ are all equal to zero, the ordinary least squares estimates of the parameters of equation (13Rb) are given by

$$(17b) \quad \hat{\Gamma}^b = (F_b' F_b)^{-1} F_b' Y$$

where Γ^b is the parameter matrix $[\Gamma_{2,7}, \Gamma_{8,12}, \gamma_1]$. Postmultiplying (17b) by r , and substituting Yr from (16b') yields

$$\hat{\Gamma}^b r = (F_b' F_b)^{-1} F_b' F_b p$$

or $\hat{\Gamma}^b r = p'$. Transposing, one gets $r' \hat{\Gamma}^b = p'$. By comparing equations (13Rb) with (13) one sees that $\hat{\Gamma}^b$ is an estimate of the parameter matrix $[AB_1 \ AB \ Z]$. Making this substitution, and also substituting the definition of p' , one gets

$$(18b') \quad r' \hat{Z} = r'$$

$$(18b'') \quad r' A \hat{B}_1 = 1$$

$$(18b''') \quad r' A \hat{B} = 0$$

Since $\hat{Z} = I_5 - \hat{A}^b$, substituting into (18b') gives $r'(I_5 - \hat{A}^b) = 0$ or $r' \hat{A}^b = r'$, which is the first condition in (11b). The second condition, $r' \gamma_{\Delta W} = 0$ is satisfied by assumption. Substituting $r' \hat{A}^b = r'$ into (18b'') and (18b''') gives conditions (12a) which, hence, are also satisfied by the ordinary least squares estimators.

III

The model of Brainard and Tobin consists of identities which imply restrictions on the coefficients of their static and dynamic behavioral equations. While Brainard and Tobin recognized these restrictions, indeed emphasized them, they did not derive them formally. Such a derivation has been accomplished in Section I of this essay. In Section II we showed how estimates of the structural coefficients of the model which satisfy the restrictions derived in Section I could be derived from ordinary (unconstrained) least squares estimates of a reduced system of equations.

REFERENCES

- W. C. Brainard and J. Tobin, "Pitfalls in Financial Model Building," *Amer. Econ. Rev. Proc.*, May 1968, 58, 99-122.
 F. E. Hohn, *Elementary Matrix Algebra*, New York 1964.

Clothing Exemptions and Sales Tax Regressivity: Note

By DAVID G. DAVIES*

In a recent article in this *Review*, Jeffrey Schaefer reports an important discovery that runs counter to currently accepted views. He finds that clothing exemptions in the New Jersey tax law actually "reduce sales tax progressivity."¹ It would be of great value if Schaefer's results could be generalized to a wider geographical area. Other state legislatures considering adopting or revising the sales tax, for example, could be spared the mistake of exempting clothing in the belief that such an exemption pushes the sales tax toward progressivity if, in fact, the opposite is true.

This comment extends Schaefer's analysis to the whole of the United States. We derive progressivity-regressivity indexes for a sales tax which excludes clothing from taxation for 1) all urban areas in the United States, and 2) all urban and rural areas combined. The empirical information for these areas comes from the most recent survey of consumer expenditures and income for the United States by the Bureau of Labor Statistics (1964 and 1966).

Following Schaefer's definitions, a sales tax is considered to be progressive if the effective rate of taxation increases as the ability to pay increases. If the effective tax rate declines as the ability to pay increases, the tax is regressive; and if the rate remains approximately constant as ability to pay changes, the tax is proportional.

An effective way to test whether a sales

tax is progressive or not is to derive the elasticity of the tax base with respect to the measure of ability to pay. We can use Schaefer's regression equation to derive the progressivity-regressivity index of alternative sales tax bases:

$$(1) \quad \sum_{j=1}^n w_j \log \bar{l}_{ij} = N \log a_i + b_{ik} \sum_{j=1}^n w_j \log \bar{y}_{jk} + \sum_{j=1}^n w_j \log e_{ijk}$$

where

\bar{l}_{ij} = mean i th sales tax base for the j th income class,

\bar{y}_{jk} = mean income of the j th income class under the k th income concept,

b_{ik} = the progressivity-regressivity index for the i th sales tax base under the k th income concept,

a_i = the constant, and

e_{ijk} = the error term for the i th sales tax base for the j th income class under the k th income concept.

If the regression coefficient, b , is greater than one, that particular sales tax base is progressive. If b is approximately equal to one, the tax is proportional; and it is regressive if b is less than one.

Although annual measured money income is a fairly generally accepted criterion of ability to pay, we shall also use Irving Fisher's notion of income and Harold Somers' and Joseph Launie's concept of net resources as alternative measures of ability to pay in deriving the progressivity-regressivity indexes for alternative tax bases. Conceptually, Fisher income is equal to consumption,² and net resources are defined as annual net income plus net worth.

² Fisher reasoned that consumption is the destruction of utility and comes closest to measuring real income. See Irving Fisher.

* Professor of economics, Duke University. I wish to express my appreciation to Professor Jay Salkin and Mr. Stanley Warner for advice and research assistance. I would also like to thank an anonymous referee for comments which improved the paper considerably.

¹ Although there has been no published empirical study of the equity effect of taxing clothing other than Schaefer's contribution, the currently accepted view is that including clothing in a sales tax base helps to make the levy regressive. See, for example, Alfred D. Buehler (p. 229), Frank Greenway, Paul Hastings, and John Smale (pp. 434-36), Niel Jacoby (p. 183), and American Federation of Labor and Congress of Industrial Organization (p. 77).

TABLE 1—PROGRESSIVITY-REGRESSIVITY INDEXES OF DIFFERENT SALES TAX BASES USING ANNUAL INCOME, FISHER INCOME, AND NET RESOURCES AS ALTERNATIVE MEASURES OF ABILITY TO PAY^a

	Urban Population			Urban Plus Rural Population		
	Annual Income	Fisher Income	Net Resources	Annual Income	Fisher Income	Net Resources
T_1 =Taxable commodities (normal tax base in food taxing states)	.87 (.04)	1.08 (.02)	1.10 (.14)	.79 (.05)	1.03 (.02)	1.04 (.14)
T_2 = T_1 minus clothing expenditures	.84 (.05)	1.04 (.02)	1.05 (.14)	.77 (.05)	1.00 (.02)	1.00 (.14)
T_3 =Taxable commodities (normal tax base in states not taxing food)	.97 (.04)	1.21 (.02)	1.24 (.14)	.88 (.04)	1.14 (.01)	1.16 (.13)
T_4 = T_3 minus clothing expenditures	.94 (.05)	1.17 (.02)	1.20 (.14)	.86 (.05)	1.11 (.01)	1.13 (.14)
T_5 =Clothing expenditures	1.07 (.03)	1.32 (.04)	1.38 (.13)	.95 (.04)	1.22 (.03)	1.27 (.12)
T_6 =Expenditures on children's (under 18) clothing	1.27 (.07)	1.58 (.05)	1.61 (.21)	1.04 (.08)	1.36 (.03)	1.37 (.18)

^a Standard errors are in parentheses.

The coefficients in Table 1 reveal that the exemption of clothing always pushes the sales tax toward regressivity, regardless of whether or not a state includes or excludes food from the tax base,³ whether or not annual measured money income, Fisher income, or net resources are used to calculate the equity coefficient, and whether or not we consider the urban population or the entire urban-rural population of the United States.

A tax base of clothing per se would yield a progressive tax for urban residents in the United States if net annual income, Fisher income, or net resources are used as the measures of ability to pay. A tax on rural plus urban expenditures for clothing is roughly proportional on net income, but progressive for Fisher income and net resources.

We also find from the coefficients in Table 1 that the progressivity-regressivity indexes of spending for children's clothing exceeds the indexes of total clothing expenditures. Moreover, a tax on children's clothing is progressive for all geographical areas and income

concepts with the exception of annual income for the combined urban-rural total where the levy is proportional. For equity purposes, exempting children's clothing from taxation, as Connecticut has done, would appear to be worse than exempting total clothing.⁴

In conclusion, it seems clear that Schaefer's important findings about New Jersey can be extended to cover both the urban and entire rural-urban population of the United States. All states considering adoption or changes in sales tax laws should be aware that exempting clothing from taxation not only erodes the tax base, but makes the levy less progressive or more regressive.

REFERENCES

- A. D. Buehler, *General Sales Taxation*, New York 1932.
- I. Fisher and H. W. Fisher, *Constructive Income Taxation*, New York 1942.
- F. Greenway, P. Hastings, and J. Smale, "Consumption Taxes" in *Public Finance*, New York 1959.

³ Currently thirty states include food while fifteen exclude it from the tax base.

⁴ The relatively high elasticity of children's clothing may be explained by the positive correlation between income and number of children.

- N. Jacoby, *Retail Sales Taxation*, Chicago 1938.
- J. M. Schaefer, "Clothing Exemptions and Sales Tax Regressivity," *Amer. Econ. Rev.*, Sept. 1969, 59, 596-99.
- H. M. Somers assisted by J. J. Launie, *The Sales Tax*, Sacramento 1964.
- American Federation of Labor and Congress of Industrial Organization, *State and Local Taxes*, Washington 1958.
- U.S. Bureau of Labor Statistics, *Consumer Expenditures and Income, Urban United States, 1960-61*, Washington 1964.
- , *Consumer Expenditures and Income, Urban and Rural United States, 1960-61*, Washington 1966.

An Economic Theory of the Second Moments of Disturbances of Behavioral Equations

By HENRI THEIL*

Applied econometricians have known for a long time that the smooth curves in economics texts which pretend to be the geometric representations of behavioral equations present an over-simplified picture of the real world. It almost never occurs that data fit a simple theoretical relation exactly, and an explicit method of interpreting and handling the deviations from such relations is therefore needed. In the early days of econometrics it was frequently assumed that observational errors were the only or at least the main cause of the occurrence of deviations, but this idea was later rejected on the ground that even when the data are perfect, deviations will continue to be found because there are very few economic agents who react exactly according to some simple mathematical law. The "neglected factors" entered into the picture; their combined influence came to be regarded as a random disturbance with zero mean and unknown variance.

It is important to realize that this disturbance variance is as much an unknown parameter of the model as the coefficients of the explanatory variables are. Typically, however, the economic theorist concentrates his efforts exclusively on these variables and their coefficients. He assumes that the decision maker whose behavior is described by the equation maximizes or minimizes a criterion function:

$$(1) \quad f(x) = f(x_1, \dots, x_n)$$

where $x = [x_1 \dots x_n]'$ is a vector of decision variables, and that the extremum is sought subject to a set of q constraints:

$$(2) \quad g_h(x) = 0 \quad h = 1, \dots, q$$

Under appropriate conditions on the form of the functions (1) and (2), he will obtain

as the solution a vector \bar{x} which expresses each decision variable uniquely in terms of the parameters that determine the functions $f(\cdot)$, $g_1(\cdot)$, \dots , $g_q(\cdot)$. Consumer demand theory is a classical example; then (1) is the utility function and (2) is the budget constraint, and the decision variables x_1, \dots, x_n are the quantities bought which are expressed by the demand functions (the behavioral equations) in terms of income and prices (the $n+1$ parameters of the budget constraint).

The disturbing thing about this approach is that it leaves no room for a random disturbance. It is undoubtedly true that there are factors which are neglected by the approach, so that discrepancies between the behavior predicted by the equation(s) and the observed behavior must be expected, but it is equally true that the economic theory in this form has nothing to say about the variance of such discrepancies. This problem has gained in importance in the last few decades since the introduction of simultaneous equation systems and "seemingly unrelated regressions," which require consideration of the disturbance variances of several equations and also of their contemporaneous covariances. When there are L such equations and when the j th contains N_j unknown coefficients, their total number is

$$(3) \quad N_1 + N_2 + \dots + N_L$$

and the total number of variances and contemporaneous covariances (taking account of the symmetry of their matrix) is

$$(4) \quad \left(\frac{1}{2}\right)L(L+1)$$

When L is moderately large, the number (4) may be of the same order of magnitude as the number (3) or even larger.

Economic theory in its conventional form focuses on the expectational part of behavioral equations by making more or less

* University of Chicago.

explicit statements on the coefficients whose number is given in (3). The question arises whether it is possible to extend this theory in such a way that it can also be brought to bear on the second moments whose number is given in (4). The objective of this article is to show that, if a set of behavioral equations refers to one single decision maker (as in the consumer demand case discussed following (2)), a plausible argument can be made for the proposition that the Hessian matrix¹ of the criterion function is the crucial determinant of the covariance matrix of the disturbances.² Basically, the idea in the case of one single decision variable is that, if the second-order derivative at the point of the extremum is close to zero, the loss incurred by a decision which deviates from the optimum to a moderate extent is very small, and the disturbance of the corresponding behavioral equation may then be expected to have a large variance.

I. The Unconstrained Case

We start with the case in which there are no constraints, q of (2) being zero, and assume that $f(\cdot)$ of (1) is to be minimized. In many economic theories, this is a smooth and well-behaved function and we shall confine ourselves to this type by assuming that it can be approximated to a sufficient degree of accuracy by a quadratic function around the minimum:

$$(5) \quad f(x) = a'x + \left(\frac{1}{2}\right)x'A x$$

where the constant term is disregarded because it is irrelevant from the viewpoint of minimization. The matrix A is the (symmetric) Hessian matrix of $f(\cdot)$ evaluated at the point of the minimum; it will be assumed to be positive definite. The decision vector which minimizes $f(\cdot)$ is then

$$(6) \quad \bar{x} = -A^{-1}a$$

Consider any decision x which differs from

\bar{x} . The associated value $f(x)$ of the criterion function will exceed $f(\bar{x})$ and the excess is the following quadratic form:

$$(7) \quad \begin{aligned} f(x) - f(\bar{x}) &= \frac{1}{2}(x - \bar{x})'A(x - \bar{x}) \\ &= \frac{1}{2}\xi'A\xi \end{aligned}$$

where

$$(8) \quad \xi = x - \bar{x} = x + A^{-1}a$$

is the decision discrepancy vector. Given that A is symmetric and positive definite, there exists a nonsingular matrix P such that

$$(9) \quad P'P = A$$

It follows from (7) that the excess can be written as

$$(10) \quad f(x) - f(\bar{x}) = \frac{1}{2}\xi'P'P\xi = \frac{1}{2}\eta'\eta$$

where

$$(11) \quad \begin{aligned} \eta &= P\xi = Px + PA^{-1}a \\ &= Px + P(P'P)^{-1}a \\ &= Px + (P')^{-1}a \end{aligned}$$

The second and third steps in (11) are based on (8) and (9), respectively.

Equation (10) shows that minimizing $f(\cdot)$ is equivalent to the minimization of the sum of squares of the elements of η . This is the simplest possible reduction of the extremum problem, and we shall therefore refer to the η elements as the *elementary decision variables*. If the decision maker definitely wants to minimize $f(\cdot)$, he should put each of these variables equal to zero. If he behaves differently, the loss which he incurs—the excess of $f(\cdot)$ over the minimum value—is equal to one-half of the squared length of η .

To introduce the stochastic element we suppose that the vector a of the linear part of $f(\cdot)$ in (5) fluctuates randomly, while the matrix A of the quadratic part remains fixed as before.³ It follows from (6) that the opti-

¹ The matrix of second-order derivatives.

² The reader who is statistically oriented will recognize the formal similarity with the asymptotic covariance matrix of maximum-likelihood estimators, which is related to the Hessian matrix of the logarithmic likelihood function.

³ The simplest way to visualize the case of a stochastic vector a and a nonstochastic matrix A is by interpreting $f(\cdot)$ as a cost function. Then all marginal cost functions are subject to additive random shocks. If the matrix A were also random, the marginal cost functions would be subject to both additive and multiplicative random shocks, which is much more difficult to handle.

mal vector \bar{x} then becomes a linear function of random variables, so that it is also random. The vector \bar{x} is then no longer feasible as a decision vector. Furthermore, when x is the (nonrandom) decision which is actually made, η of (11) is equal to the sum of the random vector $(P')^{-1}a$ and the nonrandom vector Px . In other words, the decision maker has no longer full control over his elementary decision variables. Part of each such variable is controlled, but the rest is a linear function of random variables.

The minimization of $f(\cdot)$ no longer suffices as a criterion of this stochastic situation. The criterion to be developed here rests on the assumption that the decision maker is interested in a simple mechanism to control the situation. This seems reasonably realistic, given that the decision makers involved are typically consumers or entrepreneurs who have little inclination toward formal analysis. Specifically, consider the following:

1) Given that η is random, the decision maker no longer controls this vector, but he can at least adjust the decision vector x over time so that he controls the *distribution* of η . It follows from (10) that the loss associated with any decision is equal to $\frac{1}{2}\eta'\eta$, which is a random variable. We assume that the decision maker is interested in controlling its average value, $\frac{1}{2}E(\eta'\eta)$, and this implies that the only feature of the distribution of η which is relevant is the matrix of its second-order moments.

2) Since the criterion $\frac{1}{2}E(\eta'\eta)$ concerns only the squares of the elementary decision variables, not their products, there is no gain in designing a particular cross-moment pattern. The simplest solution is then to arrange a diagonal moment matrix $E(\eta\eta')$.

3) For each of the elementary decision variables, the loss associated with a nonzero value is equal to one-half of its square. There is therefore no reason to let any such variable have a larger mean square than any other, which leads to the specification of equal diagonal elements in the moment matrix $E(\eta\eta')$. On combining this with the conclusion reached under (2) we obtain

$$(12) \quad E(\eta\eta') = \sigma^2 I$$

where σ^2 is a measure for the loss which the

decision maker is willing to incur on the average. The criterion mentioned under (1), $\frac{1}{2}E(\eta'\eta)$, is then σ^2 multiplied by one-half the number of decision variables.

It follows from (8) and (11) that the difference between the actual decision x and the theoretically optimal⁴ decision \bar{x} is $\xi = P^{-1}\eta$. This difference is the disturbance vector of the equation system which describes the decision maker's behavior with respect to the variables which he controls. Given the specification (12), we obtain the following moment matrix for this disturbance vector:

$$(13) \quad \begin{aligned} U(x - \bar{x}) &= P^{-1}E(\eta\eta')(P')^{-1} \\ &= \sigma^2(P'P)^{-1} = \sigma^2 A^{-1} \end{aligned}$$

which thus turns out to be a positive multiple of the inverse of the Hessian matrix of the criterion function.

When there is only one decision variable, this result implies a disturbance variance which is inversely proportional to the second-order derivative d^2f/dx^2 evaluated at the point of the minimum.⁵ As stated at the end of the introduction, this simply means that the disturbance variance is larger (smaller) when the loss associated with a given discrepancy $x - \bar{x}$ is smaller (larger). When we have two decision variables, the moment matrix (13) becomes

$$(14) \quad \begin{aligned} &\sigma^2 \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}^{-1} \\ &= \frac{\sigma^2}{a_{11}a_{22} - a_{12}^2} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{bmatrix} \end{aligned}$$

This indicates that, when the two decision variables do not interact in the criterion function (5) in the sense that a_{12} vanishes, the associated disturbances have a zero

⁴ Theoretically optimal in the sense of optimal under nonstochastic conditions.

⁵ When the criterion function is not quadratic, this derivative is a function of x and its value at \bar{x} is random when \bar{x} is random. One may then decide to take the expectation to obtain an approximate value. Note further that the word "variance" used here in the text is applicable only when the disturbances have zero mean. This condition is usually satisfied due to the presence of a constant term in the equation.

cross-moment. If they do interact, the cross-moment has a sign opposite to that of a_{12} . If they interact to a sufficient degree, a_{12}^2 being close to $a_{11}a_{22}$, the moments are all large in absolute value and the moment matrix is close to singularity. These results are intuitively plausible.

We conclude by noting that the model developed above treats the decision which is actually made (x) as nonstochastic and the theoretically optimal decision and its determining factors (\bar{x} and a) as random variables, whereas it is customary in the standard linear regression model to treat the dependent variable (our decision variable) as random and the explanatory variables (the determining factors) as taking fixed values. This difference is not essential, however, because the conditional distribution of $x - \bar{x}$ given \bar{x} has the same moment matrix as the conditional distribution given x .

II. The Constrained Case

Next assume that there are $q > 0$ constraints (2) subject to which the function (5) is to be minimized. We shall linearize these constraints in the same way as the criterion function is quadratized:

$$(15) \quad Bx = b$$

where the elements of B are to be interpreted as first-order partial derivatives of the functions $g_1(\cdot), \dots, g_q(\cdot)$ defined in (2), evaluated at the point of the constrained minimum. It is assumed that B has full row rank (i.e., rank q), so that the possibility of inconsistent or linearly dependent constraints is excluded.

One way of solving this more general problem is by using the constraints (15) to eliminate q of the x 's and then proceeding along the lines of the previous section. However, a more elegant extension can be formulated on the basis of the Lagrangian function:

$$(16) \quad \begin{aligned} F(x, \lambda) &= f(x) + \lambda'(Bx - b) \\ &= [a' - b'] \begin{bmatrix} x \\ \lambda \end{bmatrix} \\ &\quad + \frac{1}{2} [x' \quad \lambda'] \begin{bmatrix} A & B' \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} \end{aligned}$$

The new decision vector is $[x' \lambda']'$ rather than x , the coefficient vector of the linear part is now $[a' - b']$,⁶ and the original Hessian matrix is bordered by q rows and q columns. It is readily seen that the extension of the moment matrix result (13) to the present case amounts to σ^2 times the inverse of the bordered Hessian matrix, and that the leading sub-matrix (corresponding to A in the bordered Hessian before inversion) of this inverse gives the moment matrix of the disturbance vector corresponding to x . Application of the standard rules for partitioned inversion gives:

$$(17) \quad \mathcal{V}(x - \bar{x}) = \sigma^2 [A^{-1} - A^{-1}B'(BA^{-1}B')^{-1}BA^{-1}]$$

The matrix (17) is zero in the special case of a square nonsingular B . This is as it should be, because the decision vector can then be solved from the constraint (15). Note further that premultiplication of the moment matrix (17) by B gives a zero matrix, which is also in agreement with (15).

III. Application to Consumer Demand Theory

Consider a consumer whose objective is to maximize a utility function $u(x)$ subject to the budget constraint $p'x = m$, where p and x are n -element vectors of prices and quantities, respectively, and m is total expenditure (or income). It is assumed that p and m are fixed and given from the consumer's point of view. This amounts to the following specification of the more general model described above: $f(\cdot) = -u(\cdot)$, $B = p'$, $b = m$. Write U for the Hessian matrix of $u(\cdot)$ in the constrained maximum, so that A is specified as $-U$. The moment matrix (17) then becomes

$$(18) \quad -\sigma^2 \left(U^{-1} - \frac{1}{p'U^{-1}p} U^{-1}pp'U^{-1} \right)$$

which is a positive semi-definite matrix of order $n \times n$ and rank $n - 1$ when U is negative definite. The singularity of the moment matrix is due to the budget constraint $p'x = m$.

⁶ The form of this coefficient vector indicates that the vector b can be allowed to fluctuate randomly over time, just as a .

The matrix (18) is precisely the same as the moment matrix which I derived earlier on the basis of *ad hoc* considerations.⁷ These considerations do add to the plausibility of the general model, however, and it is therefore worthwhile to summarize them here briefly without going into mathematical details:

(1) When equating $-f(x)$ of (5) to $u(x)$, we obtain

$$-a_i + \frac{1}{2} \sum_j u_{ij} x_j$$

for the marginal utility of the i th commodity, a_i and u_{ij} being the i th element of a and the (i, j) th element of U , respectively. The assumption of a random vector a implies that the marginal utilities are subject to additive random shocks. Assume that they have zero expectation. The main problem, to be considered under (2), is the specification of the $n \times n$ covariance matrix of these shocks.

(2) Consider the case $u_{ij}=0$, so that the i th and j th marginal utilities do not depend on x_j and x_i , respectively. This indicates that

the two commodities satisfy unrelated wants, so that it seems reasonable to assume that the random shocks of their marginal utilities are uncorrelated. Next consider the case of a *negative* u_{ij} . The i th marginal utility then decreases with x_j , which may be interpreted in the sense that the two commodities satisfy similar wants. When the marginal utilities are subject to random fluctuations, those corresponding to similar wants will usually be of the same sign and will therefore be *positively* correlated. Third, consider a commodity with a large negative u_{ii} , so that its marginal utility is unstable in the sense of being sensitive to small changes in q_i . It is then plausible that it is also unstable in the sense that the random shocks of its marginal utility have a large variance. All three cases ($u_{ij}=0$, $u_{ij}<0$, $-u_{ii}$ large) are covered when it is assumed that the covariance matrix of the random shocks is a negative multiple of the Hessian matrix U , and this leads directly to the matrix (18).

REFERENCE

H. Theil, *Economics and Information Theory*, Amsterdam; Chicago 1967, pp. 228-33.

See Theil, pp. 228-33.

A Neglected Social Cost of a Voluntary Military

By THOMAS E. BORCHERDING*

It would appear without exception that economists believe that a voluntary military is preferable to conscription.¹ It is my purpose to demonstrate that this institutional preference is questionable on purely a priori grounds. A *potentially* important welfare cost may arise under voluntarism from the monopsonistic behavior of the defense establishment² as a purchaser of enlisted personnel. To analyze this possibility it will be necessary to develop a terse and simple model of choice in the "market" for enlisted personnel and to apply it to the institutions of conscription and voluntarism.

Assume that the process of political exchange is efficient to a degree that the military's demand function for enlisted men is a close approximation to the social marginal value schedule of this input. Further, the supply curve of this resource is taken to be an approximation of the value of its social alternatives. Given the usual assumptions for a specified level of employment, a measure of net social benefit can be derived by measuring the area lying between these schedules (*DD* and *SS* in the figure following). Institutional considerations aside, the Paretian level of employment is shown at *OA*, where social benefits are maximized. By imposing a coercive subsidy on this factor, the draft leads the military to "pur-

chase" an excessive amount, *OB*, at the draft supply price \bar{p} . The welfare loss associated with this allocation is in excess of the area of the triangle $\alpha\beta\gamma$, since certain other costs do not show up on this diagram. These latter arise from 1) excessive training costs associated with high personnel turnovers; 2) the evasion and avoidance efforts of potential conscriptees attempting to escape the draft; and 3) the cost associated with drafting some individuals whose opportunity costs exceed those of others who are exempted.

Distributional considerations aside, the proponents of voluntarism hold that it would be optimal since under this regime the military is forced to pay the market supply price of this factor instead of the subsidized price, \bar{p} . It will be demonstrated that this is not strictly correct since it neglects the possible deadweight burden associated with monopsonistic purchase of volunteers. Such behavior may at first sound unreasonable; but the potential for it is implicit in the assumptions made by voluntarists in their condemnation of the draft together with the evidence concerning the supply function of that factor.

Since the voluntarists hold that the draft leads to an excessive purchase of enlisted men, they implicitly assume that the military treats the budgetary cost of that input under conscription as the actual cost. Does not consistency require that this assumption hold under voluntarism as well?³ Further, since the empirical evidence indicates that the supply function is upward sloping (see

* The author, assistant professor at the University of Washington, thanks John S. McGee, Yoram Barzel, Lowell R. Bassett, and Allan Hynes for their many helpful suggestions. He also extends his appreciation to the Reim Foundation, whose summer grant for 1969 financed this research as well as other work still in process.

¹ Mark Pauly and Thomas Willet are the most recent and best known of this group. Their paper also surveys past positions on these matters. Other papers in the same volume are useful to read on the subject of efficiency and military procurement.

² By defense establishment, I mean not only the military and the Department of Defense, but Congress which acts as an intermediary between the voters and taxpayers and the DOD.

³ Consider, for example, the possibility that the military actually does treat the supply price of enlisted personnel as their cost. It would follow, then, that under voluntarism the optimal quantity would be purchased. It would also be true that under conscription the correct numbers would be drafted. Except for distributional considerations, and perhaps certain other inefficiencies of conscription already discussed, there would be little to choose between the two institutions.

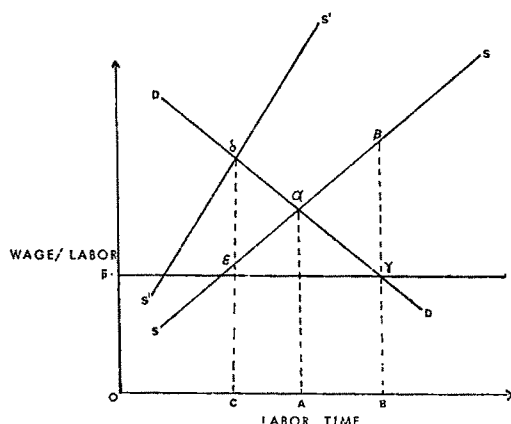


FIGURE 1

Stuart Altman and Alan Fechter, and Walter Oi), monopsonistic purchase is a distinct possibility unless wage discrimination is possible.⁴

This latter phenomenon would bring the marginal budgetary cost in line with the supply price, but it would appear unlikely for two reasons. First is the usual information problem: the determination of individual reservation prices is neither easy nor costless. Second, and more important, is the political constraint: equal wages for identically defined jobs are now the law of the land. Given the unlikelihood of wage discrimination, the relevant incremental factor-cost function to the military would not be the supply schedule, SS , but its marginal, $S'S'$.⁵ The resulting monopsonistic allocation, OC , occasions a welfare loss equal in area to the triangle $\alpha\delta\epsilon$.

Although we have some information about the supply function, without some knowledge

of the demand schedule and, possibly, some notion of the other costs associated with the draft, it is impossible to specify which inefficiency is more damaging. A rough idea about the magnitude of AB and CA might be helpful. For example, if we could know that AB , the overextension of purchase under conscription, exceeds CA , the underutilization of voluntarism, then a simple assumption that DD and SS are linear (or convex to the horizontal axis) over this range would lend great support to the voluntary military scheme. If, instead, CA were the larger, actual calculation of the welfare triangles $\alpha\beta\gamma$ and $\alpha\delta\epsilon$ would be required, as well as some estimate of the other costs of conscription.

Whether this counter-example points out a serious social cost or is a mere pathological construction based on a doubtful premise would appear an interesting area for further exploration. Such a project would require not only a fair amount of empirical study but, more importantly, a thoughtful examination of our current model of public choice.

REFERENCES

- S. H. Altman, "The Structure of Nursing Education and Its Impact on Supply," in H. E. Klarman, ed., *Empirical Studies in Health Economics*, Baltimore 1970.
- and A. E. Fechter, "The Supply of Military Personnel in the Absence of a Draft," *Amer. Econ. Rev. Proc.*, May 1967, 57, 19–31.
- R. L. Bish and P. D. O'Donoghue, "A Neglected Issue in Public Goods Theory: The Monopsony Problem," *J. Polit. Econ.*, forthcoming.
- J. M. Buchanan, *Cost and Choice*, Chicago 1969.
- W. Oi, "The Economic Cost of the Draft," *Amer. Econ. Rev. Proc.*, May 1967, 57, 39–62.
- M. V. Pauly and T. D. Willett, "Efficiency in Military Manpower Procurement," in J. C. Miller III, ed., *Why the Draft? The Case for the Volunteer Army*, New York 1968.

⁴ In Altman's work the notion of monopsony purchase by a public agency is employed to explain the existence of vacancies. I have recently discovered that Robert Bish and Patrick O'Donoghue have also articulated the problem as does James Buchanan (pp. 89–92).

⁵ Empirical evidence suggests that the elasticity of supply is between 1.2 and 1.8 (see Altman and Fechter, and Oi). Thus $S'S'$ must lie from 50 to 85 percent higher than SS .

On the Extension of Input-Output Analysis to Account for Environmental Externalities

By A. O. CONVERSE*

Robert Ayres and Allen Kneese have extended input-output analysis to include the flow of wastes to the environment and the recycle flow of wastes to the production sectors. This was done by adding two sectors, consumption and environment. All "output" from the consumption sector is either recycled or sent to the environmental sector. In addition to this material, the environmental sector also receives flows directly from the various production sectors.

In the Ayres and Kneese model, the flow of waste from the k th production sector to the environment is given by $C_{k0}X_k$ where C_{k0} equals the fraction that comes from the k th sector and X_k is the total mass of residuals discharged to the environment.¹ Assume that the output of the non-waste commodity from the j th sector increases. This would cause an increase in raw materials required and hence, an increase in X_0 . Let Δ be the amount of this increase. The increase in the flow of wastes from the k th sector would be $C_{k0}\Delta$ whether the production in that sector were changed or not. Furthermore, the increase from the j th sector would be $C_{j0}\Delta$ rather than proportional due to the change in the amount of the j th commodity. If such a model were used to distribute a pollution tax, it would be an unfair distribution since the tax would not be directly related to the amount of residuals being discharged to the environment by the particular sector.

In the usual production sectors of the Leontief input-output analysis, the inputs are set by the output from that sector. However, in the case of waste inputs to the en-

vironmental sector, the inputs are set by the outputs of *other* sectors. This rather fundamental relationship is absent from the Ayres and Kneese formulation.

I. Waste from the Production Sectors

The model should be revised so that the flow of waste residuals from the k th production sector to the environment is given by $C_{k0}X_k$, where X_k is the flow of the non-waste commodity from the k th sector and C_{k0} is the ratio of the waste residuals to the non-waste commodity in the k th sector. This would overcome the above objections.

In making the above modification one begins to account for the multiproduct nature of production, which is after all, the basic cause of pollution from the production sectors.² The above form can account for only one type of waste residual from a given production sector. The model could easily be extended as follows to account for the full range of products.

Let

- (1) $v_{ij}U_j$ = amount of i th commodity produced or consumed by the j th activity

where:

U_j = extent of the j th activity, as measured by any *one* of its inputs or outputs, or perhaps by the amount of fixed capital invested.³

v_{ij} = amount of the i th product or commodity used or produced per unit

* Associate professor of engineering, Thayer School of Engineering, Dartmouth College.

¹ X_0 is equal numerically to the total amount of raw materials withdrawn from the environment, inventories being neglected.

² This aspect is important in other applications of input-output analyses. For example, Alan Manne pointed out the importance of this in his analysis of the petroleum industry.

³ Perhaps some function of the fixed capital invested would be best as it could account for non-constant returns to scale.

activity in the j th sector, positive for products and negative for inputs.⁴

With this extension one can account for the various types of waste residuals and not merely the total amount. This is obviously very important when trying to evaluate pollution effects. As a matter of fact, the whole treatment effort is based on the assumption that it is desirable to increase the total amount of waste residuals discharged to the environment in order to change the composition.⁵ One possible use of an extended input-output simulation would be to test this assumption.

II. Waste from Consumption Activities

The consumption inputs, Y_i , are set either by the planner or the market. During consumption they are not really consumed but merely transformed into wastes.

$$(2) \quad W_k = \sum_i \alpha_{ki} Y_i$$

Whereas the above relationship can account for different types of wastes, Ayres and Kneese lump them all into one item. The increased detail will give rise to greater computational effort but it may well be worth it. For certain applications the detail of the breakdown could be quite modest, e.g., products that are discharged into the atmosphere and those that are discharged into the surface waters. These wastes, W_k , are subject to further processing,⁶ hence one must expand equation (1) to

$$(3) \quad X_i = \sum_j \nu_{ij} U_j + \sum_k \gamma_{ik} W_k$$

⁴ This corresponds to a stoichiometric coefficient in a chemical reaction. The use of such notations was suggested to the author by Mr. Peter Brooks, University of Queensland.

⁵ The total amount is increased because the pollution control treatment requires increased use of resources.

⁶ If there really were no further processing, then setting $\gamma_{kk}=1$ and $\gamma_{ik}=0$ for all $i \neq k$, would formally convert all "wastes" into "commodities."

We now distinguish between three types of commodities:

Goods for consumption, $Y_i = X_i$, assuming that demands are met.

Materials withdrawn from the environment = X_i , when $X_i < 0$.

Materials discharged to the environment = $X_i - Y_i$, when $X_i - Y_i > 0$.

Conservation of mass requires that

$$(4) \quad \sum X_i - Y_i = 0$$

all i corresponding to material commodities

It should be noted that the above equality holds only when the flows, X and Y , are expressed in units of mass.

III. Summary

The modification of input-output analysis presented by Ayres and Kneese does not correctly account for the individual waste residues from the various production sectors. A modest change that overcomes this objection is presented. Further modifications that would allow one to account for the various types of waste residues from both production and consumption activities are presented. The need for such detail is caused by the specific activities of the various residues (CO_2 is significantly different from CO). It is noted that pollution treatment while changing the composition of the waste residues does increase the total amount of them. Hence any analysis that considers only the total amount will be unable to evaluate pollution control measures.

REFERENCES

- R. U. Ayres and A. V. Kneese, "Production, Consumption, and Externalities," *Amer. Econ. Rev.*, June 1969, 59, 282-97.
 A. S. Manne, "A Linear Programming Model of the U.S. Petroleum Refining Industry," *Econometrica*, Jan. 1958, 26, 67-106.

Mishan on the Gains from Trade: Comment

By MEL KRAUSS AND DAVID M. WINCH*

In a noted article, Harry G. Johnson developed an analytical apparatus for measuring the welfare gains from international trade in terms of "the goods that could be extracted from the economy in the free trade situation without making the country worse off than it was under protection—some variant of the Hicksian compensating variation (p. 329).¹ In a general equilibrium model, Johnson first divides the welfare effect of a change in commercial policy—the removal of an autarkic tariff on imports, for example—into gross changes in consumers' and producers' surplus, arriving at a *net* welfare effect defined in terms of consumers' surplus; then subdivides this net welfare effect into two separate and distinct components, the first reflecting the *consumption cost* of the tariff; the second, the tariff's cost in terms of deviation from the optimal pattern of production, i.e., the tariff's *production cost*.

This approach to the welfare theory of tariffs has been questioned by E. J. Mishan in a recent article in this *Review* (pp. 1280–82). Though Mishan appears to accept Johnson's breakdown of the welfare effect of trade into consumption and production cost components, he is not willing to go along with Johnson's expression of the gain from trade (or loss from a tariff) as a net gain of consumer's surplus on the grounds that the "... division of the welfare gain from free trade into gains (or losses) of consumer's surpluses offset by losses (or gains) of producer's surpluses is ... arbitrary" (Mishan, p. 1281), and for that reason—and judging from Mishan's argument, for that reason alone—erroneous. While one could hardly disagree that Johnson's expression of the

gains from trade as a net gain in consumer's surplus is arbitrary, and indeed must be arbitrary given the impossibility of distinguishing buyer and seller in an exchange model without fiat money, the authors cannot conclude in the same breath that being arbitrary is the same thing as being erroneous.² Indeed if one were to extend Mishan's logic beyond the case in point, the whole of the literature that utilizes the concept "numeraire" would have to be discarded on this account alone.³

I

The analysis is best begun by seeking a "correct" interpretation of the concept "surplus." Certainly, the terms consumers' surplus and producers' surplus are misnomers; surplus arises from exchange not from production and consumption, and a change in surplus normally derives from a change in the conditions of exchange, not from the conditions of production or consumption. In a general equilibrium context, surplus arises from exchange of factors for products by households or of products for factors by firms. Though, by necessity, there must be at least two parties to an exchange, the surplus arises from the exchange, and not from either side of the bargain. One can isolate consumers' (or buyers') surplus only by pegging the factor side of the household's exchange to the numeraire by assumptions

² It should be noted that in this paper the authors address themselves solely to Mishan's Appended Note (pp. 1280–82). Since Mishan himself freely uses the terms "producers' surplus" and "consumers' surplus" in his note (in seeming contradiction to his recommendation in the main text of his article that the term producers' surplus be struck from the economist's vocabulary), the present authors are reluctant to discontinue this practice.

³ Since the choice of a numeraire in exchange models without fiat money is necessarily arbitrary, and to be arbitrary is to be wrong according to Mishan, one cannot escape the conclusion that all theory developed from models using numeraire is erroneous.

* Associate professor and professor of economics, respectively, McMaster University, Hamilton, Ontario.

¹ It should be noted that the compensating variation technique is obviously superior to the classical approach of relating the gains from trade to changes in an index expressing one or another concept of the terms of trade.

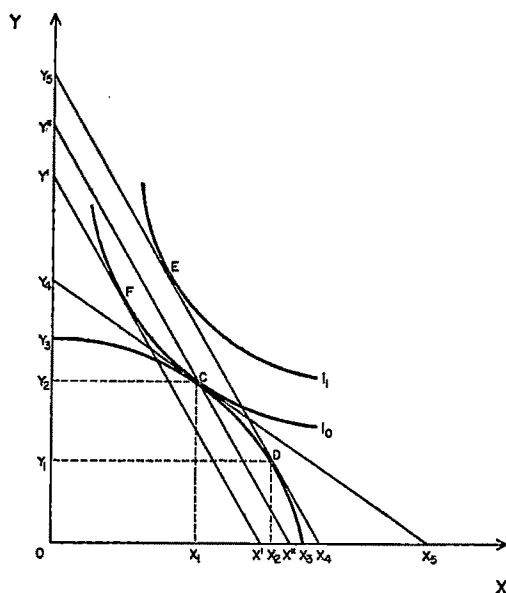


FIGURE 1

about constant income or wage rate. Similarly, producers' (or sellers') surplus is meaningful in isolation only if one pegs all other variables to the numeraire by constant price assumptions. It is quite arbitrary whether a change in the factor to product price ratio is held to yield the household a consumer's surplus (or buyer's rent) as buyer of the product, or a producer's sur-

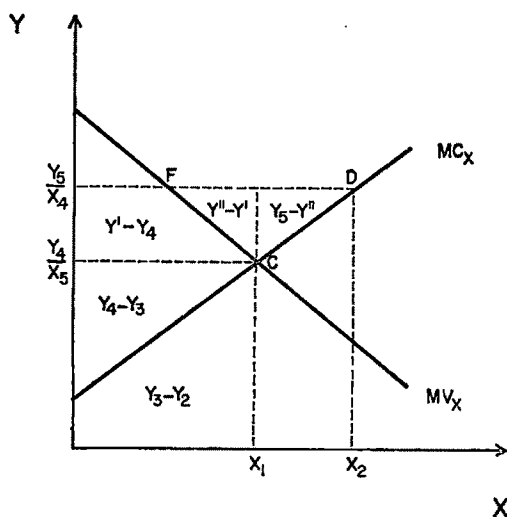


FIGURE 2

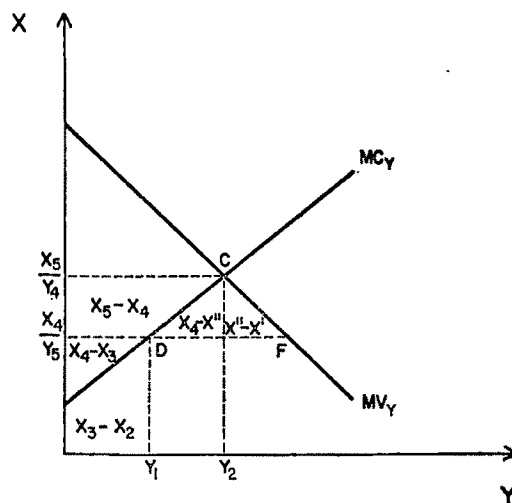


FIGURE 3

plus (or seller's rent) as a seller of a factor. Similarly, it can yield the firm a surplus (profit) as buyer of factor or seller of product. Since in a competitive equilibrium, there is no pure profit (firms' surplus), all surpluses can be attributed to either factor sellers or product buyers depending on the choice of numeraire.

II

The transformation curve x_3y_3 of Figure 1, derived from the familiar Edgeworth-Bowley efficiency box (not shown), is combined with a set of comparable community indifference curves I_0I_1 , to yield in Figure 2, a supply (marginal opportunity cost) curve of the export commodity x and a compensated demand curve for x (marginal valuation curve), y being the numeraire; in Figure 3, a supply curve of the import commodity y with a compensated demand curve for y , x being the numeraire. According to the numeraire chosen (an arbitrary choice and one that makes no real difference to the result), the movement from self-sufficiency to free trade (from point C to D) can be shown alternatively as a movement from the price ratio Y_4/X_5 to Y_5/X_4 in Figure 2, or X_5/Y_4 to X_4/Y_5 in Figure 3 (i.e., in terms of numeraire x or numeraire y). In Figure 2 where y is the numeraire, producer's surplus rises and consumers' surplus falls. This is because the total payment to factors expressed in y rises

while the purchasing power of y in buying x falls (the price of x in terms of y rises). Similarly, the movement from autarky to free trade can be shown in Figure 3 where x is the numeraire. Here the total payment to factors expressed in x falls while the purchasing power of x in buying y rises (the price of y in terms of x falls). The sum of the two changes ($Y_5 - Y'$) or ($X_4 - X'$) is the total welfare effect of moving from I_0 at point F to I_1 at point E in Figure 1. Whether this is represented as a net gain in producers' surplus (Figure 2) using y as the numeraire, or a net gain in consumers' surplus (Figure 3), using x as the numeraire, is as arbitrary as the choice of the numeraire itself. But being arbitrary doesn't make it erroneous. To paraphrase Marshall, it is a matter of indifference which numeraire is used to express a given surplus.

There can be no doubt that Johnson was aware of the numeraire problem inherent in general equilibrium models without fiat money, and thus to the potential misuse of his technique by others in double counting the welfare gain from trade as excess consumers' surplus on x plus excess producers' surplus on y .⁴ The two measures are neither separable in the sense that they have real (i.e., utility or welfare) significance in the absence of a numeraire, nor are they addable in the sense that their sum accurately portrays the gain that accrues to the community from the opportunity to trade. Being alternatives, either is sufficient for the pur-

pose of measuring the change in welfare. They are alternative depictions of the same thing, not separate components of a single whole.

A final note relates to Mishan's suggestion that the welfare gain from trade is *best* measured as a single compensating variation at the new international price ratio—as Y_5Y' or X_4X' —and interpreted as an exact measure of the gains for the community as a whole in moving from the consumption possibilities presented by X_3Y_3 to the new consumption possibilities presented by X_4Y_5 , with no mention being made of separating this measure into consumption cost and production cost components. While Mishan's overall measure is correct,⁵ indeed it is identical with Johnson's, the latter's technique of separation has the substantial advantage of highlighting the sources of the welfare gain, i.e., the gain from substituting lower cost for higher cost goods in consumption, and the gain from diverting resources from direct higher cost to indirect lower cost production of goods that can be imported from the world market. The failure to make use of this empirically relevant and theoretically illuminating analytical distinction is simply to throw out useful analysis for the sake of misplaced methodological purism.

REFERENCES

- J. Bhagwati and H. G. Johnson, "Notes on Some Controversies in the Theory of International Trade," *Econ. J.*, Mar. 1960, 70., 74-93.
- H. G. Johnson, "The Cost of Protection and the Scientific Tariff," *J. Polit. Econ.*, Aug. 1960, 68, 327-45.
- E. J. Mishan, "What is Producer's Surplus?," *Amer. Econ. Rev.*, Dec. 1968, 58, 1269-82.

⁴ This is evident from Johnson's discussion of the changes in the terms of trade that might be induced by moving from a restricted trade to free trade position, and the effect this would have on the measure of the gains from freer trade, i.e., the quantity of goods that could be extracted from the economy, leaving it as well off under free trade as it was with the tariff. On p. 330, Johnson clearly states that the terms of trade gain or loss will differ according to whether the compensating variation is effectuated with exportables or importables. If exportables are used, i.e., if x is the numeraire, the extraction results in a terms of trade gain for the extracting country; if, on the other hand, importables are used, the result is a terms of trade loss. Cognizance of this problem clearly implies cognizance of the general problem of numeraire in general equilibrium models without fiat money, and thus the dangers of double counting. Also, see Jagdish Bhagwati and Johnson, where three alternative measures of the gains from trade are developed, each using a different numeraire.

⁵ Mishan's triangle EHG of his Figure b (p. 1280) is equivalent to our triangle DCF in Figure 3, and the sum of Johnson's two triangles GHI and DEF in his Figure 2 (p. 331). The difference between Mishan and the present authors (and Johnson) is that Mishan contends that the identification of triangle EHG as a net gain of consumers' surplus over a loss of producers' surplus is erroneous. It should be noted that Mishan's position on this issue appears is no way related to his general condemnation of the concept "producers' surplus" expressed in the main text of his article.

Mishan on the Gains from Trade: Reply

By E. J. MISHAN*

Apart from some conventional sniping from behind improvised footnotes, the counterattack mounted by David Winch and Mel Krauss has the apparent aim of recapturing for Harry Johnson the right to continue to use producers' surplus in his analysis of the welfare effects of tariff protection. Since they do not appear to dispute the main thesis developed in the text of my paper (that producers' surplus is either a misnomer or a conceptual error), their gallantry in attempting to secure a special dispensation for the use of producers' surplus in the gains-from-trade case also involves a certain awkwardness; one that is, not surprisingly, reflected in their arguments. Broadly speaking, their tactics are first to churn up a terminological smokescreen so as to obscure traditional notions of consumers' surplus and rents, and then, while the innocent reader is blinking at the resulting swirl, to produce for him concepts guaranteed to be slippery enough to meet any contingency.

I

By adapting Johnson's tariff argument to a special autarky-free trade situation, I affirmed that under special conditions—zero income elasticity of importables, Pareto-comparability of community indifference curves, and the adoption of the free trade terms of trade to measure welfare gains—there is a clear mathematical equivalence of the welfare gain of removing a tariff when measured in terms of a vertical (or horizontal) distance on the transformation-indifference curve construction and when measured in terms of an area bounded by the derived marginal curves. The identification by Winch and Krauss of these ways of measuring the net gain of moving from autarky to free trade does not, therefore, add anything to the argument in my Appended Note. Contrary to what they seem to allege, I did

not repudiate Johnson's breakdown of the welfare effects because they were arbitrary with respect to *measure* (being measurable, in the two-good case, in terms either of x or y) but simply because they were arbitrary with respect to *concepts*; by which I refer to Johnson's interpretation of the net community gain as an excess of the gain of consumers' surplus over the loss of producers' surplus or, equally invalid, as an excess of the gain of producers' surplus (in terms of the other good) over the loss of consumers' surplus.¹

The issue between Winch and Krauss and myself is, therefore, simply this: that they believe the *net* gain from trade can be properly interpreted as an excess of a gain of consumers' over a loss of producers' surplus, or the reverse, and I do not.

In their Section I on the subject of surpluses, "consumers' surplus" and "producers' surplus" are asserted to be misnomers, on the grounds that "surplus arises from exchange, not from the conditions of production or consumption." Consistency of interpretation, however, requires that a measure of surplus be regarded as a measure of the *change* in a person's welfare irrespective of how it is brought about—whether an effective constraint is removed, whether a change in one or more prices favor him, whether the conditions of his work or his environment improve or deteriorate, or whether his wife leaves him. As has already been pointed out elsewhere (J. R. Hicks), the traditional division into a consumer's surplus (in which, narrowly conceived, a per-

¹ In their fn. 2, the authors write: "Since Mishan himself freely uses the terms producers' surplus and consumers' surplus in his note (in seeming contradiction to his recommendation in the main text . . .) the present authors are reluctant to discontinue this practice." As a *non sequitur* this is also ingenuous. It should have been obvious that, in paraphrasing and extending Johnson's argument, prior to registering my dissent, there was some provisional convenience in employing his own terminology.

* London School of Economics.

son's welfare is affected by changes in one or more product prices, all other product and factor prices constant) and rent (in which, narrowly conceived, a person's welfare is affected by changes in one or more factor prices, other factor prices and all product prices remaining constant) has occasional convenience in welfare analysis—although it should be borne in mind, always, that any gain or loss resulting from a complex alteration of both factor and product prices can, in principle, be measured in terms of income, or in terms of one or more of the constant-priced goods.

Their further pronouncements on this subject add nothing to this, while subtracting a great deal in the way of clarity. Thus producers' surplus is smuggled in again—in a rather casual way, so as not to give offence. For they talk of a 'producer's surplus' (or seller's rent) as accruing to the seller of a factor. From which one might reasonably infer that they would like a licence to use producer's surplus while referring to the rent of a factor. Reluctantly one sympathizes with their cunning, since a vindication of Johnson's use of producers' surplus looks to be an impossible task without injecting occasional doses of ambiguity into the argument. We are, therefore, further informed that a change in the factor-product price ratio can also yield the firm a surplus, which is profit, though in competitive equilibrium there is no firms' surplus, since any surplus must be attributed to factors or consumers. But which it is to be depends, according to the authors, on the choice of numeraire.

Given our traditional definitions, however, it depends on what occurs. If a firm disposes of its profits by lowering the price of its products, *ceteris paribus*, it generates a consumers' surplus. If, on the other hand, it does so by raising the price of its factors, it generates an increase in rents. If, however, we are dealing with the economy as a whole, and the only information we have, in the two-good model, is that the price of one good has risen and that of the other has fallen (or one risen in terms of the other), then we cannot, in general, say either that consumers' surplus has increased or de-

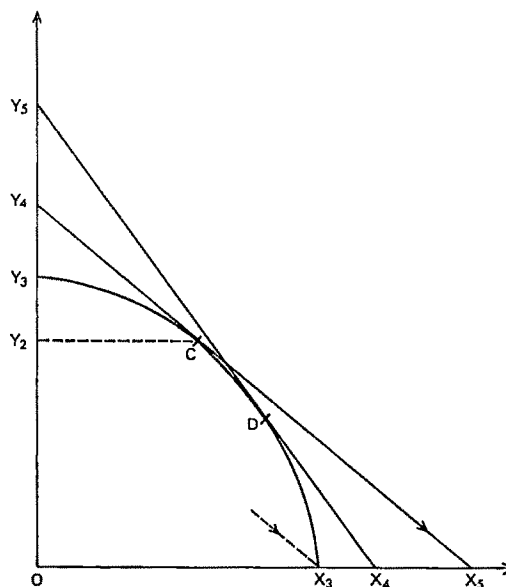


FIGURE 1'

creased, or that rents have increased or decreased. (We can say nothing about a producers' surplus which—unless it is to be a synonym for Knightian profit, in which case it is zero—does not exist.) For all that, the authors are determined to deliver to us a producers' and consumers' surplus cut out of the gains from trade. Let us see what surprises they come up with.

II

Now it transpires that the crux of their argument, for retaining the use both of a producers' and a consumers' surplus in measuring the welfare gain from trade, turns only on relative price-changes, and on *product* prices at that. It can be summarized with reference to my Figure 1' (which includes only the features of their Figure 1 that are necessary to this part of their argument).²

² Their Figures 2 and 3 are used to depict their conclusions, which conclusions depend only on the fact of a change in relative product prices. Nor is the movement from I_0 to I_1 (measured in terms of Y_5X_4) a part of their demonstration that there is both a producers' surplus and a consumers' surplus in terms of one good or the other. This measurement of gain from I_0 to I_1 is, as I argued, a net gain of community welfare and can, as I asserted, be measured in terms of either x or y .

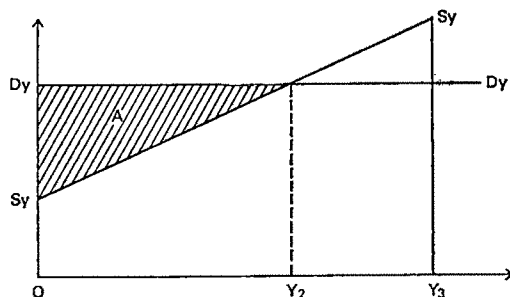


FIGURE 3'

Consider, then, the movement from C to D , along with the accompanying change in the product terms of trade—these being indicated by the shift in the tangential straight lines Y_4-X_5 to Y_5-X_4 ; and let us forget about producers' surplus inasmuch as the authors' argument, in the first paragraph of their Section II, becomes strictly in terms of factor rents. Have rents changed? Has consumers' surplus changed? Yes, they answer. For a) if we take product y as the numeraire, total rents (i.e., the rents of the two factors A and B) rise from Y_4 to Y_5 whereas total consumers' surplus falls because the price of x rises in terms of the numeraire y . On the other hand, according to the authors, b) if we take product x as the numeraire, total rents fall from X_5 to X_4 , whereas consumers' surplus rises because the price of y has risen in terms of the numeraire x . This apparently is the logic that produces an excess of producers' surplus over the loss of consumers' surplus for case a), and the reverse, an excess of gain of consumers' surplus over the loss of "producers' surplus" for case b). So it all depends, after all, on the numeraire.

The reader will observe in passing: i) that these relative changes occur simply in moving from C to D , regardless of whether such movement is from autarky to trade, or from trade to autarky, or the result of a change in tastes or of a change in the distribution of the product; ii) that, at the relevant terms of trade, batch C is valued at Y_4 or X_5 while batch D is valued at Y_5 or X_4 . Therefore, when the movement from C to D is accompanied by a rise in the price of x relative to

that of y , the community *as a whole*—whether they are regarded as consumers or producers³—looks better off in terms of y and looks worse off in terms of x . From this change, it is not possible to infer an opposition of welfare as between consumers as a whole and earners as a whole; and iii) that if x is A -intensive then, as the Stolper-Samuelson theorem affirms (for two-input production functions homogeneous of degree one), the owners of factor A are *absolutely* better off in the movement from C to D , whereas the owners of factor B are *absolutely* worse off. But no means are devised in these constructs for determining whether, *on balance*, total real rents have increased or not. Consequently it is just not possible to establish that the real increase of welfare which (given Pareto-comparability of I_0 and I_1) occurs in the movement from C , the autarkic position, to D , the free trade position, is the resultant of a real increase in rents that is greater than a real decline in consumers' surplus, or the reverse.

Indeed, as I pointed out, the welfare gain is represented on the original diagram as a measure only (at the international terms of trade) of moving from community indifference curve I_0 to community indifference curve I_1 . Neither the welfare of factors or of firms can be distinguished or isolated from the original indifference-transformation construction, since they are not there to begin with.

What then are we to make of the plausible-looking areas in Winch and Krauss's Figures 2 and 3?

III

A useful clue is provided if we begin with a Robinson Crusoe economy, in which point C is chosen on the transformation curve⁴ of my Figure 1', with the tangential line Y_4-X_5

³ In the sense that if the terms of trade consistent with point D could be maintained for any amount of x exchanged for y , the community could end up with OY_5 of y , whereas with the terms of trade consistent with point C the community could end up with only OY_4 of y .

⁴ No reason need be given for its convexity, nor for how it comes into being. To Crusoe, the locus of alternative combinations of x and y is to be accepted as a datum.

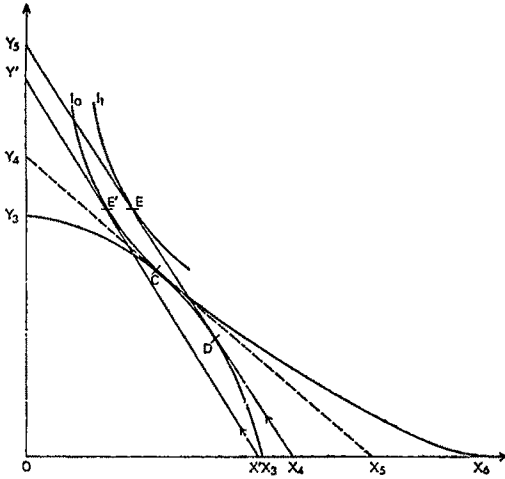


FIGURE 1''

representing Crusoe's indifference curve. Taking the first derivative of the transformation function with respect to y , we construct Figure 3' (comparable with Figure 3 drawn by Winch and Krauss).

How do we interpret the shaded triangle A ? The schedule $S_y S_y$, the derivative of the transformation curve with respect to y , begins from a position in which Crusoe holds only x and shows the increasing amounts of x that have to be surrendered for each additional unit of y . The height of the horizontal line $D_y D_y$ represents the unchanged marginal value to Crusoe (in terms of x) of these additional increments of y . The area of the triangle A , therefore, measures the welfare gain of Crusoe's taking OY_2 of y on these terms.⁵ If we have to put a label to the welfare gain, we should call it a buyer's or a consumer's surplus in deference to the traditional commonsense definition that harks back to Marshall—being the difference between what he (Crusoe) is willing to give up for OY_2 of y , and what he has to give up for it. And this

⁵ If Crusoe begins with OX_3 of x and can transform x into y at a constant rate equal to his constant marginal valuation, he gains nothing from exchanging y for x . It is just because he need give up less of x than this for the initial units of y that enables him to increase his welfare, an increase that can be measured as the horizontal (or vertical) distance between the straight line indifference curve passing through X_3 and the parallel indifference curve passing through X_5 .

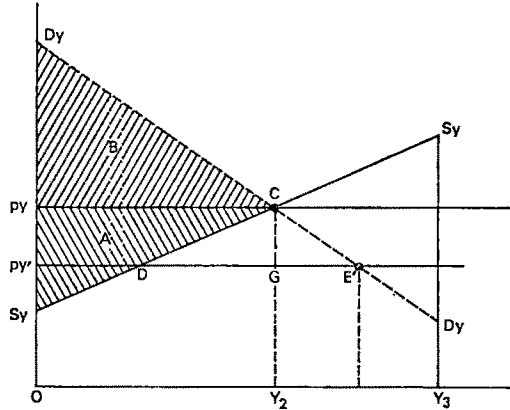


FIGURE 3''

area which is more suggestively regarded as a consumer's surplus—and *not* a producers' surplus as Winch and Krauss would have it—corresponds in my Figure 1' to the horizontal distance $X_5 - X_3$. This is the maximum amount of x that Crusoe would pay in order to have the terms of trade $Y_4 X_5$ since they enable him to move from X_3 to C .⁶

The reader will readily appreciate that no difference is made to this analysis if the indifference curve touching C took on the familiar concave shape as shown by I_0 in my Figure 1'' and, in consequence, the $D_y D_y$ curve (the marginal valuation of y in terms of x) were downward sloping as shown in Figure 3''. Crusoe's gain in welfare, still regarded as a consumer's surplus, would then be the area of the two shaded triangles A and B , a total area that corresponds to the distance $X_6 - X_3$ in Figure 1''.

Let Crusoe now be faced with world terms of trade $Y_5 X_4$ in Figure 1''. He does the best for himself by producing the combination D and exporting as much of x as is necessary to enable him to consume combination E on the I_1 curve. The gain in moving from the consumption of C on the I_0 curve to the consumption of E on the I_1 curve can be measured either in terms of x or of y , and,

⁶ This welfare gain can also be measured, at the same terms of trade, in terms of y . Alternatively, we could measure a different welfare gain (in terms of x or of y) of being presented with these same terms of trade if, instead, we suppose Crusoe to begin at Y_3 , having only y and no x .

moreover, at either the old terms of trade or at the new. If the gain from trade is measured at the new international terms of trade, and in terms of x , the movement from I_0 to I_1 can be measured as a compensating variation $X_1 - X'$; i.e., as the maximum amount of X that Crusoe is willing to sacrifice in order to have the privilege of trading at the new international terms of trade.

If we let the "income" elasticity of demand for y be zero, so that E' on I_0 is on a line through E that is parallel with the x -axis of Figure 1'', then $D_y D_y$ in Figure 3'' is the compensated demand curve. The compensating-variation measure of gain from a lower price of y in terms of x , p_y' , is accurately depicted as equal to the triangle DCE' .

The interpretation of this gain from trade is also straightforward, and has no reference to a producers' surplus, or to rents of factors. Crusoe will continue to acquire y from domestic sources until point D is reached on his $S_y S_y$ curve. Beyond that point, it is cheaper for him to buy y abroad at a fixed price p_y' (the amount of x he gives up being readily accepted as exports). And he continues to buy y at this international price p_y' until he reaches point E' , beyond which his marginal valuation of y falls below this p_y' price.

This area of gain, DCE' , can be split into two triangles: the area of the first triangle DCG being that part of the gain from being able to buy some part, DG , of Crusoe's original consumption of y at an international price, p_y' , that is lower than the domestic price, p_y . The second triangle GCE' is the additional gain arising from the extra quantity, GE' of y , that Crusoe will consume as a result of the lower international price p_y' .

The identification by Johnson of the gain represented by the first triangle, DCG in Figure 3'', as "the increase in the value of production," and that of the second triangle, GCE' in Figure 3'', as "the reduction in the cost of consumption," does, perhaps, make some sort of sense, for what the distinction is worth. For the area of the first triangle depends on the shape of the marginal transformation curve $S_y S_y$ and can, therefore, be associated with the conditions of production,

while the area of the second triangle depends upon the shape of the compensated demand curve $D_y D_y$ and can, therefore, be associated with the conditions of consumption.

What *cannot*, however, be inferred from the derived triangle DCE' in Figure 3''—whether we remain with Robinson Crusoe or whether, instead, we have Figure 1'' refer to the community—is that it is the resultant of a gain of consumers' surplus over a loss of producers' surplus. For, as we have seen—regardless of whether we obstinately think of producers' surplus in terms of profits or of rents to factors—there is no real loss of producers' surplus in moving from the production of C to that of D . There is only the compensating variation measure of community gain in moving from I_0 to I_1 ,⁷ a gain that is represented by the triangle DCE' in Figure 3''.

IV

Finally, a word on the alleged empirical usefulness of Johnson's technique.⁸ Although the analysis did not lack in sophistication, a cursory reading at least gives the impression that its value is more heuristic than empirical. Allowing that a measurement of gains (or losses) of tariff-reductions proceeds by a consideration of importables only, calculation of the net gain in the two-good case would require a supply curve and a compensated demand curve of x in terms of y .

⁷ Their "excess producers' surplus"—triangles $Y'' - Y'$ and $Y_8 - Y''$ in their Figure 2—is also equal to a compensating variation, in term of y , of moving from the consumption of C on I_0 and that of E on I_1 , this area corresponding to the vertical distance $Y_8 - Y'$ in their Figure 1.

⁸ It should be noted, however, that their alleged "substantial advantage of highlighting the sources of welfare gains, i.e., the gain from substituting lower cost for higher cost goods in consumption, and the gain from diverting resources from direct higher cost to indirect lower cost production of goods that can be imported from the world market," which can be translated into Johnson's division of the net gain into a "reduction in the cost of consumption" and an "increase in the value of production," (that correspond to triangles DCG and GCE' in my Figure 3'') is not something I take issue with. The issue, to repeat, is the interpretation of the net gain, measured by the combined area of these two triangles, as one resulting from an excess in the gain of consumers' surplus over the loss of a producers' surplus.

In the real world estimates of demand and supply curves are in terms of money prices, with money prices of other goods constant. But, first of all, once the analysis is extended to a number of goods, it is no longer reasonable to assume upward-sloping supply curves even if the model is restricted to two factors. Assuming production functions are homogeneous of degree one (or less), at least, it is no longer possible to infer unambiguously that an increase in the demand for a good will raise its long-run equilibrium supply curve either absolutely or relatively.⁹

Again the estimates of the kind of demand functions that are required to calculate the gains or losses arising from a reduction of the tariffs on a number of importables are not easy. If x_1, x_2, \dots, x_n are importables that

⁹ If there are n -goods, x_1, x_2, \dots, x_n , ranked thus in order of increasing A -intensity, a shift of demand, from say, x_4 to x_5 , raises the price of factor A (on the common assumption of some inelasticity in its supply) relative to B (though raising it by less the larger the number of A -using goods having some factor substitutability). This rise in the relative price of A raises the supply price of x_n relative to that of $x_{n-1} \dots$ relative to that of x_5 relative to that of $x_4 \dots$ relative to that of x_1 —assuming, always, the A -intensity ranking remains unaltered. Given a constant stock and velocity of money, we can say nothing of the money price of x_5 , however, without further specification. We can deduce for certain only that the money price of x_n will rise. On the other hand, a shift of demand from good x_8 to x_5 lowers the price of A relative to B , and lowers the price of x_n relative to that of $x_{n-1} \dots$ relative to that of $x_8 \dots$ relative to that of $x_5 \dots$ relative to that of x_1 . Again, without further specification, we can say nothing of the change in the money supply price of x_5 . We can deduce only that the money supply price of x_n will fall.

are, in some degree, substitutes, the gain (or loss) from a simultaneous removal of their tariffs is not to be measured by adding together the apparently relevant areas under the more usual *ceteris paribus* demand curves for each of the x_1, x_2, \dots, x_n importables. Following J. R. Hicks' analysis of 1956, the total gain is to be calculated by taking the tariff-removals in sequence; adding the gain from removing the x_1 tariff alone, the tariffs on the remaining $(n-1)$ goods unchanged, to the gain from removing the x_2 tariff alone when the x_1 tariff is already removed but the tariffs remaining on the other $(n-2)$ goods, and so on until the x_n tariff.

Demand curves estimated for these conditions are not, however, readily available. And although I am inclined to agree with Johnson that, for Western countries at least, the gain from removing tariffs is not likely to be very significant in terms of *GNP*, such conclusions depend much more on rough guesses than on any careful application of the apparatus he skillfully designed in his 1960 paper.

REFERENCES

- J. R. Hicks, *A Revision of Demand Theory*, Oxford 1956.
- H. G. Johnson, "The Cost of Protection and the Scientific Tariff," *J. Polit. Econ.*, Aug. 1960, 68, 327-45.
- E. J. Mishan, "Rent as a Measure of Welfare Change," *Amer. Econ. Rev.*, June 1959, 49, 386-95.

Profit Constrained Revenue Maximization: Note

By RICHARD ROSENBERG*

In recent issues of this *Review*, the profit constrained, revenue maximizing oligopoly model introduced by William J. Baumol has been both expanded and applied. (See articles by Robert Haveman and Gilbert DeBartolo (1968, 1970), C. J. Hawkins, Milton Z. Kafoglis, Kafoglis, and Robert C. Bushnell, and David J. Smyth.) The attention paid to this model can be interpreted as a symptom of the general feeling of dissatisfaction with regard to the lack of an acceptable theoretical explanation for much of oligopoly behavior. While Baumol has made an important advance by introducing a two-dimensional objective function for oligopoly firms, he has, unfortunately, also specified a particular form for the objective function which turns out to be untenable.

Unlike the use of constraints in defining an environment or in determining the subset of the choice set which is attainable, the use of constraints in an objective function has clear implications for both the method of arriving at the decision maker's ordering and for his behavior. To postulate that firms seek to maximize sales revenue, subject to a profit constraint, implies that firms order various outcomes (each outcome is a combination of a certain level of profits and a certain level of sales revenue) in a lexicographic manner as illustrated in Figure 1. Each of the lines in the figure is a behavior line rather than an indifference curve. This implies that the firm orders possible outcomes in the following way: 1) for any two outcomes, both of which have profits below the constraint, the outcome with the larger profit is preferred (*B* is preferred to *A*) regardless of the associated levels of sales revenue. If both have the same level of profit, the outcome with the larger sales revenue is preferred (*C* is preferred to *B*); 2) for any two outcomes, both of which have profits equal to or greater than the

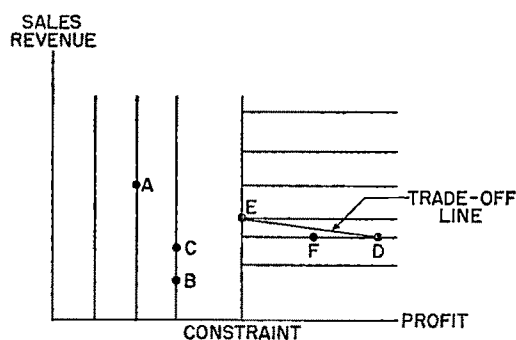


FIGURE 1

constraint, the outcome with the greater sales revenue is preferred, (*E* is preferred to *D*). If both have the same level of sales revenue, the outcome with the larger profit is preferred (*D* is preferred to *F*).

Moreover, since the model assumes that dollars of profit can always be converted into additional sales revenue through promotional activities, if the attainable set of outcomes includes one outcome for which profits exceed the constraint, then the attainable set will also include an outcome at which profits are equal to the constraint and sales revenue will be increased. In the figure, if the firm can attain a position such as *D* by profit maximizing behavior, it can and will move along a trade off line such as *DE* until it reaches the profit constraint at *E*. The firm makes this trade no matter how small the increase in revenue relative to the decrease in profit.

In conventional terminology from the theory of demand, the firm is thus assumed to have a Marginal Rate of Substitution of sales revenue for profit (the amount of profit it is willing to give up in order to receive an extra unit of sales revenue) which is infinite so long as profit exceeds the constraint, and which is always equal to zero so long as profit is below the constraint.¹ While

* Assistant professor of economics, Pennsylvania State University.

¹ This implication was recognized as long ago as 1965 by Armen A. Alchian.

such an ordering is conceptually possible, there does not seem to be any economic rationale for supposing that such a strong and unstable preference pattern should exist. While it may prove fruitful to postulate that firms are not motivated solely by the desire for greater profits, the constrained maximization approach is not a satisfactory method for embodying this notion.

REFERENCES

- A. A. Alchian, "The Basis of Some Recent Advances in the Theory of the Management of the Firm," *J. Ind. Econ.*, Nov. 1965, 14, 30-41.
- W. J. Baumol, *Business Behavior, Value and Growth*, rev. ed., New York 1967.
- R. Haveman and G. DeBartolo, "The Revenue Maximization Oligopoly Model: Comment," *Amer. Econ. Rev.*, Dec. 1968, 58, 1355-58.
- , "The Revenue Maximization Oligopoly Model: Reply," *Amer. Econ. Rev.*, June 1970, 60, 433-34.
- C. J. Hawkins, "The Revenue Maximization Oligopoly Model: Comment," *Amer. Econ. Rev.*, June 1970, 60, 429-32.
- M. Z. Kafoglis, "Output of the Restrained Firm," *Amer. Econ. Rev.*, Sept. 1969, 59, 583-89.
- and R. C. Bushnell, "The Revenue Maximization Oligopoly Model: Comment," *Amer. Econ. Rev.*, June 1970, 60, 427-28.
- D. J. Smyth, "Sales Maximization and Managerial Effort: Note," *Amer. Econ. Rev.*, Sept 1969, 59, 633-34.

Behavior of the Firm Under Regulatory Constraint: Note

By ISRAEL PRESSMAN AND ARTHUR CAROL*

A recent article by Akira Takayama discusses an earlier paper by Harvey Averch and Leland L. Johnson on fair rate of return regulation of public utilities. Although Takayama (p. 255) agrees with Averch and Johnson's general conclusions "that a firm will tend to increase its investment with the introduction of an active constraint" on its rate of return, he criticizes the A-J argument as being "confusing, ambiguous, and in error." Takayama then attempts a clarification, and presents a new formulation which leads to the A-J result quoted above.

This comment will discuss several of Takayama's criticisms in addition to showing that the so-called "A-J Effect" cannot be derived from the basic assumptions made by both Averch and Johnson and Takayama.¹ We will show that the very assumptions used to prove the A-J Effect, by defining the region of λ , require an assumption that the A-J Effect exists in the first place.

I. The Model

Consider a monopoly employing two inputs, capital (x_1) and labor (x_2) to produce a single homogeneous output (z). The firm faces a production function

$$z = z(x_1, x_2),$$

defined on $x_1 \geq 0, x_2 \geq 0$, having positive first-order derivatives and satisfying $z(0, x_2) = z(x_1, 0) = 0$. The price (p) is related to the output by the inverse of the demand function, i.e.,

$$(1) \quad p = p(z)$$

* The authors are, respectively, assistant professor of operations research at the Polytechnic Institute of Brooklyn and associate professor of economics at the University of Hawaii. Pressman is also a member of the Management Sciences Division of the American Telephone and Telegraph Company. The authors are indebted to the referees and all those who made helpful comments on earlier drafts of this paper. The views expressed here are solely our own.

¹ Both Averch and Johnson and Takayama define these assumptions as a concave revenue function and that s_1 , the maximum allowed rate of return, is strictly greater than r_1 , the cost of capital.

The profit is defined by

$$(2) \quad \pi = pz - r_1x_1 - r_2x_2$$

with r_1 and r_2 the factor prices presumed constant. The revenue, pz , is assumed to be concave. The regulatory constraint on the rate of return is given as²

$$(3) \quad \frac{pz - r_2x_2}{x_1} \leq s_1$$

The problem then is to maximize (2) subject to (3). The Lagrangian $L(x_1, x_2, \lambda)$ is formed and the Kuhn-Tucker necessary conditions for a maximum at $\bar{x}_1, \bar{x}_2, \bar{\lambda}$ are given.³

II. Analytical Aspects

Comments on Takayama's Criticisms

Consider first the question of whether λ , the Lagrange multiplier, has a range $0 \leq \lambda < 1$. Averch and Johnson (p. 1055) prove that $0 < \lambda < 1$ by noting three things: first, "that $\lambda = 1$ if, and only if, $r_1 = s_1$;" second, since $s_1 > r_1$, $\lambda \neq 1$; and finally, λ varies continuously. Takayama criticizes this proof as to the question of continuity of λ with s_1 . He then indicates conditions under which λ will be continuous.⁴

To derive the conditions of continuity, Takayama writes an equation for λ as

$$(4) \quad \lambda = \frac{G_K - r}{G_K - s}$$

² See Averch and Johnson, pp. 1054-55, for discussion of the assumptions leading to this formulation of the constraint.

³ Averch and Johnson, p. 1055.

⁴ We note that Averch and Johnson fail to prove the "if" part of their statement, i.e., $r_1 = s_1$ implies $\lambda = 1$. In fact, no analytic proof known to the authors has been given in the literature to support this assumption. An alternate proof of the range of the Lagrange multiplier is presented by William Baumol and Alvin Klevorick. Their proof, however, depends on the same basic assumptions required by Averch and Johnson, i.e., a concave revenue function and $s_1 > r_1$.

assuming $G_K - s \neq 0$,⁵ and noting that the r in equation (4) is the r_1 in the A-J notation.

We note that since $s > r$, the value of λ will depend on whether G_K is greater than both s and r , less than both s and r , or greater than r but less than s . If G_K is greater than both s and r , then $\lambda > 1$. If G_K is less than both s and r , then $0 < \lambda < 1$. These results are entirely possible if the constraint is an equality and Lagrange Multiplier optimization techniques are employed. Thus equation (4) above gives no clue as to the value of λ .⁶

Finally, Takayama criticizes Averch and Johnson for assuming that the marginal-revenue-product-of-capital curve (*MRPK*) does not shift. He claims (p. 257) that this assumption "is not true in general," and "involves an error of confusing the movement along the curve with the shift of the curve." If, however, this criticism is valid and the *MRPK* does actually shift with changes in labor and capital, then the assumption " G_K is a continuous function of L^* and K^* " (p. 259), (which is made to define the conditions under which λ^* is continuous) is not necessarily valid. In addition we know of no evidence mathematically or empirically that " L^* and K^* are continuous functions of s ."

Comments on the A-J Effect

A major objection to both Averch and Johnson and Takayama is the assumption that so-called A-J Effect can be derived from the basic assumptions. Actually, the two assumptions⁷ needed to prove the A-J Effect (by defining the region of λ) require the existence of this effect. Consider the A-J formulation of the problem given above, i.e.,

⁵ We note that G_K is equivalent to Averch and Johnson's

$$\left(p + z \frac{dp}{dz}\right) \frac{\partial z}{\partial x_1}$$

In addition, Takayama uses y, K, L, r, w , and s where Averch and Johnson use z, x_1, x_2, r_1, r_2 , and s_1 .

⁶ Takayama's further contention that $G_K - s = 0$ implies $s = r$ is correct; however, equation (4) in no way includes the assumptions $s > r$, nor does it define a value of λ when $G_K - s = 0$.

⁷ a) concave revenue function, and b) $s_1 > r_1$.

$$\max \pi = pz - r_1x_1 - r_2x_2$$

$$(5) \quad \text{s.t. } pz - s_1x_1 - r_2x_2 \leq 0$$

$$x_1, x_2 \geq 0$$

The essential discussion of A-J is when the constraint is operative,⁸ i.e.,

$$(6) \quad pz - s_1x_1 - r_2x_2 = 0$$

Thus, a solution to the problem (by inspection) is to increase x_1 (capital) to a value greater than the unconstrained value x_1^0 , say to x_1^* . This would assure a reduction of the unconstrained rate of return s_0 to the regulated value s_1 . That this is the solution to problem (5) has also been shown geometrically by Eugene Zajac. Therefore, we wish to investigate only the case $x_1^* > x_1^0$. Writing the Lagrange function for problem (5) we have

$$(7) \quad L = (1 - \lambda)pz + (\lambda s_1 - r_1)x_1 - (1 - \lambda)x_2r_2$$

Taking partials with respect to x_1, x_2 , and λ we get

$$(8) \quad \alpha G_{x_1}^* + \lambda s_1 - r_1 = 0$$

$$(9) \quad \alpha G_{x_2}^* - \alpha r_2 = 0$$

$$(10) \quad pz - s_1x_1^* - r_2x_2^* = 0$$

where $\alpha = 1 - \lambda$. Since we require that $r_1 < s_1$ we have

$$(11) \quad r_1 = \alpha G_{x_1}^* + \lambda s_1 < s_1$$

or

$$(12) \quad \alpha G_{x_1}^* < (1 - \lambda)s_1 = \alpha s_1$$

Thus

$$(13) \quad G_{x_1}^* < s_1 \text{ implies } \alpha > 0$$

Although equation (13) allows for $G_{x_1}^* < 0$, the possibility of having both $G_{x_1}^* < 0$ and $\alpha < 0$ does not occur since $\alpha < 0$ implies $\lambda > 1$ which then implies, from equation (11), that $r_1 > s_1$, a nonallowable condition. If we assume that $G_{x_1}^* < 0$, then $r_1 > G_{x_1}^*$ by defini-

tion. Then, since $r_2 = G_{x_2}^*$, either $\left(p + z \frac{dp}{dz}\right)$

⁸ Takayama, p. 256, fn. 8.

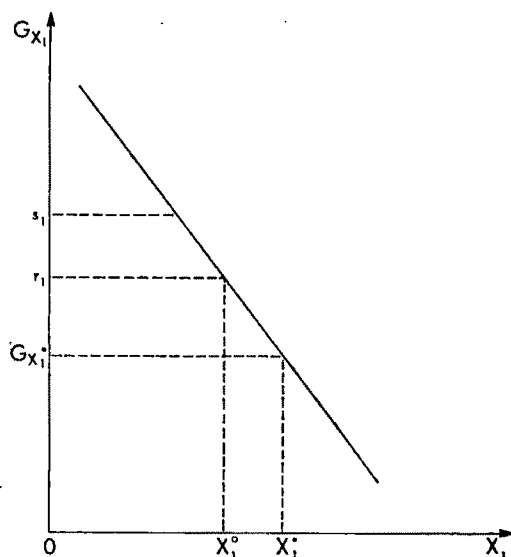


FIGURE 1

and $\partial z/\partial x_2$ are negative, or $\partial z/\partial x_1 < 0$. Thus the optimum operating point occurs where one of the marginal products is negative, i.e., in the region of economic inefficiency. In addition, if Averch and Johnson allowed for $G_{x_1}^* < 0$, then

$$r_1/r_2 > \frac{G_{x_1}^*}{G_{x_2}^*} = - \frac{dx_2}{dx_1}$$

and the A-J Effect can be seen without any discussion of λ 's. Now from equation (8) we have

$$(14) \quad r_1 = \alpha G_{x_1}^* + \lambda s_1 = G_{x_1}^* + \lambda (s_1 - G_{x_1}^*)$$

Thus, from equations (13) and (14)

$$r_1 > G_{x_1}^* \text{ if, and only if, } \alpha > 0, \lambda > 0$$

and

$$(15) \quad r_1 < G_{x_1}^* \text{ if, and only if, } \alpha > 0, \lambda < 0$$

However, the assumption $r_1 < s_1$ alone does not indicate the region of λ . Imposing the second assumption, i.e., concave revenue functions, we assume that G_{x_1} and G_{x_2} are

continuous functions of x_1 and x_2 , respectively, and that G_{x_1} and G_{x_2} can be represented by downward sloping curves as in Figure 1. The solution to the unconstrained profit maximization problem is given by

$$(16) \quad r_1 = G_{x_1}^0$$

with optimum solution x_1^0 . For the constrained problem, equation (9) yields $G_{x_2}^* = r_2$, the same as for the unconstrained problem. Thus, from our previous discussion where we assumed that $x_1^* > x_1^0$ is the solution, we see that this assumption leads to

$$(17) \quad G_{x_1}^* < r_1 < s_1$$

From equation (17), (13), and (15) we would then conclude that $\alpha > 0$ and $\lambda > 0$ which implies $0 \leq \lambda < 1$. If, however, we assume no relationship between x_1^* and x_1^0 , then we have no indication as to the position of $G_{x_1}^*$ relative to r_1 and s_1 and thus do not know the region of λ . Thus, a proof for the range of λ between 0 and 1 which depends on the continuity of G_{x_1} with x_1 must already assume a relationship between x_1^* and x_1^0 .

We see, therefore, that the A-J Effect is valid only if one accepts the two basic assumptions and one assumes either that $x_1^* > x_1^0$, or that $0 < \lambda < 1$. To prove the A-J Effect without either of these latter assumptions is impossible since the continuity of $\lambda = 0$ to $\lambda > 0$ cannot be shown in general.

REFERENCES

- H. Averch and L. L. Johnson, "Behavior of the Firm Under Regulatory Constraint," *Amer. Econ. Rev.*, Dec. 1962, 52, 1052-69.
- W. J. Baumol and A. K. Klevorick, "Input Choices and Rate of Return Regulation: Overview of the Discussion," *Bell J. Econ.*, autumn 1970, 1, 162-90.
- A. Takayama, "Behavior of the Firm Under Regulatory Constraint," *Amer. Econ. Rev.*, June 1969, 59, 255-60.
- E. E. Zajac, "A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model," *Amer. Econ. Rev.*, Mar. 1970, 60, 117-25.

Spectral Analysis and the Detection of Lead-Lag Relations

By JOHN C. HAUSE*

The difficulties of determining timing relationships between aggregate economic series by direct inspection of the series, by the cross correlation function, or by the formulation of explicit models are notorious. Empirical efforts to establish such relationships frequently fail or are highly inconclusive. In the search for techniques to deal with these problems, some economists have concluded that spectral and cross spectral analysis can avoid these difficulties and can provide direct and relevant information about leads and/or lags between pairs of economic time-series. A casual reading of studies such as those by Vittorio Bonomo and Charles Schotta, and T. J. Sargent might lead to the (incorrect) belief that it is generally possible to infer leads and lags by a simple transformation of the phase from cross spectral estimates.¹ This misunderstanding is found in a recent paper by Thomas Cargill that attempts to test the hypothesis that wage changes lag significantly behind price changes solely by an examination of cross spectral estimates.

This paper demonstrates that the interpretation of phase statistics depends critically upon the model (or class of models) that one assumes governs the relationship between a pair of time-series. Only under extremely restrictive assumptions is it valid to interpret phase information in the time domain in the way that Cargill proposes.

The discussion falls under two headings. The first deals with deterministic linear systems, distributed lags, and the interpretation of phase. The second section considers stochastic linear systems, the cross spectrum, leads and lags, and some final conclusions.

* Associate professor of economics at the University of Minnesota, and currently a research fellow at the National Bureau of Economic Research.

¹ These papers use the procedure described later in the text for converting phase shifts to the time domain without discussing the conditions under which this transformation is theoretically meaningful.

The paper does not discuss statistical estimation.

I. Deterministic Linear Systems, Distributed Lags, and the Interpretation of Phase²

Much of the terminology of spectral analysis, *including the concepts of lead and lag*, originated in the analysis of deterministic systems (and their mathematical representations) by engineers. The engineering definitions of lead and lag are not closely related to the intuitive notion that most economists associate with these terms. The engineering concepts were originally used to describe phase relations between variables in the frequency domain. The engineering term corresponding most closely to the economist's notion of a lag in the time domain is "pure delay" (or simply delay.) Confusion over this terminology has doubtlessly been the main source of misunderstanding by economists of the interpretation of phase statistics. A brief analysis of deterministic linear models and the corresponding terminology clarifies these issues and provides the necessary background for interpreting spectral phase statistics.

Consider the following simple model of a stationary dynamic linear system $y(t) = Lx(t)$, in which $x(t)$ and $y(t)$ are variables (functions of time) representing input and output, respectively. A linear operator, L , characterizes the way in which the input is transformed into output.³ It is generally

² Many detailed discussions of deterministic linear systems are available. An excellent, compact discussion is included in chapter 2 of G. M. Jenkins and D. G. Watts. For a more expanded treatment, see W. Kaplan.

³ In the context of this model, an operator is linear if it transforms simultaneously applied inputs into output according to the relation

$$L(a_1x_1(t) + a_2x_2(t)) = a_1Lx_1(t) + a_2Lx_2(t)$$

where a_1 and a_2 are constants. The operator is "stationary" (time-invariant) if $Lx(t) = y(t)$ implies $Lx(t+t_0) = y(t+t_0)$, i.e., a translation of the input function in time by t_0 units implies the output is translated by the same amount of time. See Jenkins and Watts

possible to represent a linear dynamic input-output relationship as a distributed lag:

$$(1) \quad y(t) = \int_0^{\infty} w(\tau)x(t - \tau)d\tau,$$

where $w(\tau)$ is a distributed lag function.⁴

An important characteristic of this class of models (for stable systems) is that if the input variable is a sinusoid of a specified frequency f , e.g., if $x(t) = \cos(2\pi ft)$, the output will be a sinusoid of the same frequency of the form

$$y(t) = G(f) \cos(2\pi ft + \alpha(f))$$

once the "transients" die out, where $\alpha(f)$ and $G(f)$ are real functions of f ,⁵ $\alpha(f)$ is the *phase shift* and is usually assigned a value between $-\pi$ and $+\pi$ (radians) by convention.⁶ According to this convention, one says that at a given frequency f output

(p. 36) or Kaplan for further discussion of stationary linear systems. Linear difference equations and the most widely used distributed lags in economics are examples of such systems.

⁴ For discrete time models used in most econometric work, this expression is usually written in the form $y_t = \sum_{k=0}^{\infty} w_k x_{t-k}$, where at least one w_k differs from zero.

⁵ $G(f)$ is the *gain* of the system, a nonnegative function that measures the amplitude of the output sinusoid if the amplitude of the input sinusoid is one.

⁶ There are two conventions in use that should be distinguished, since the terminology and conventions discussed in this section are also relevant for cross spectral statistics. For the one described in the text (in which $-\pi \leq \alpha < \pi$; $G(f) \geq 0$), $G(f)$ is called the (cross) amplitude as well as the gain. The alternative convention restricts the phase range so that $-\pi/2 \leq \alpha < \pi/2$, and allows $G(f)$ to be positive or negative. (The source of these alternatives stems from the square root and the arctangent being multivalued functions.) The definitions of lead or lag (which depend on the conventional sign of the phase angle) are the same for these alternatives if the gain is positive according to both conventions; but they are reversed if the gain is negative according to the second convention.

Unfortunately, these conventions are not discussed adequately in many treatments of spectral analysis, including Jenkins and Watts, and C. W. J. Granger and M. Hatanaka. A computer program for estimating the cross spectral statistics based on such references might well adopt a bastardized (and for some purposes, misleading) convention with $F(f) \geq 0$, and $-\pi/2 \leq \alpha < \pi/2$. It would be highly advisable for economists using cross spectral computer programs to make certain which convention is being adopted in the computation of phase and gain statistics.

leads input if α is positive, and that output lags input if α is negative. Thus the original definitions of lead and lag in the analysis of this class of models were simply a way of describing the shift in phase of the output sinusoid (*on the frequency domain, not the time domain*) and depend on a purely conventional way of measuring the size and sign of this phase shift.

The distributed lag formulation of this input-output model makes it clear that the output sinusoid is not in general obtained by a simple displacement of the input sinusoid in time by the phase angle α even though a graph of the input and output functions might lead one's intuition to this erroneous conclusion. The output is the superposition of the previous values of the input, with weights given by the distributed lag function $w(\tau)$. In general, $\alpha(f)$ will vary with frequency of the input even though the distributed lag itself is invariant to changes in frequency.

The concept of delay corresponds to the shift in time which many economists have in mind when they discuss time lags between economic variables. A (linear) delay between an input and output may be expressed by the equation

$$y(t) = ax(t - t_0)$$

where a is a constant and t_0 is the length of the time delay. This is the only deterministic linear system for which there is a simple correspondence between the phase lead or lag and delay in real time. If the input is $x(t) = \cos(2\pi ft)$ and the system is a pure delay operator, the output

$$y(t) = a \cos(2\pi f[t - t_0])$$

Hence the phase $\alpha(f) = -2\pi ft_0$.⁷ Dividing α by the frequency (in radians) of the input

⁷ Even if the pure delay model is appropriate, the convention that determines α , as discussed in fn. 6, must be handled with care. For the convention where $-\pi \leq \alpha < \pi$, if f is greater than $1/2t_0$, $-2\pi ft_0$ will fall outside the conventional range. In this case, an integral multiple of π must be added to the conventionally measured phase to obtain the total phase angle delay $2\pi ft_0$. This problem is discussed further in Granger and Hatanaka.

sinusoid gives the length of the delay, since $\alpha/2\pi f = -t_0$. For all other linear input-output systems, it is incorrect to interpret $\alpha/2\pi f$ as if there is a pure delay between input and output.

One further important characteristic of deterministic linear systems useful in the next section is cited without proof. If the system is represented in distributed lag form, the Fourier transform of the distributed lag of $w(t)$ is

$$\int_{-\infty}^{\infty} e^{-i2\pi ft} w(t) dt = G(f) e^{i\alpha(f)}$$

where $\alpha(f)$ and $G(f)$ are the phase and gain as defined above. This formula shows that the gain and phase of a linear system can be obtained directly from the Fourier transform of the distributed lag of the system.

II. Stochastic Linear Systems, the Cross Spectrum, and Leads and Lags⁸

It can be shown that the cross spectrum of two stationary random processes, x and y , is identical to the Fourier transform of the cross correlation function of x and y , and that the cross spectrum can be written in the form

$$A_{xy}(f) e^{i\beta_{xy}(f)}$$

where $A_{xy}(f)$ and $\beta_{xy}(f)$ are real functions of f called the cross amplitude spectrum and phase spectrum, respectively.

Suppose that some stochastic elements are introduced into the simple linear input-output model discussed in Section I. Using distributed lag notation, let output

$$(2) \quad y(t) = \int_0^{\infty} w(\tau) x(t - \tau) d\tau + u(t),$$

where $u(t)$ is an additive random component. Assume $u(t)$ is statistically independent of $x(t)$. The cross spectrum of input and output of this system is $C_{xy}(f) = W(f) C_{xx}(f)$, i.e., the cross spectrum is the product of the Fourier transform of the distributed lag

function and the spectrum of x . The spectrum must be a real nonnegative function. The preceding section pointed out that $W(f)$ can be written in the form $G(f) e^{i\alpha(f)}$. Thus for this linear model

$$A_{xy}(f) = [G(f) C_{xx}(f)] \quad \text{and} \quad \beta_{xy}(f) = \alpha(f)$$

The last equation states that the phase angle of the (deterministic) distributed lag in this case is identical to the phase of the cross spectrum of input and output. The result is not surprising, since spectral analysis essentially amounts to a frequency decomposition of time series. For each frequency, the phase relationship of input and output is precisely that determined by the distributed lag.

The discussion in Section I emphasized that the phase α of a deterministic system can be interpreted as a shift in time only if the distributed lag is a pure delay. The calculations and conclusions in Cargill's paper are based on the assumption that phase has been determined in a pure delay system. The variable τ is defined by the relation $\alpha(f)/2\pi f$, and τ is discussed as if it measures a pure delay at that frequency. A simple, but revealing illustration in the Appendix demonstrates how misleading this procedure can be if the distributed lag is not a pure delay.

The emphasis upon the analysis of pairs of variables in which one variable can be regarded as a distributed lag of the other stems from the important role this model has played in thinking about the time relationships between certain aggregate economic variables. E. Malinvaud (p. 473) has given particular emphasis to the role of distributed lags in econometric investigations and has suggested why they are more plausible than pure delays. If one turns to linear models which are not simple input-output systems, the interpretation of phase (and τ) becomes even more complicated. For example, the pair of variables used for computing cross spectral statistics might themselves be distributed lags of a third variable. The nature of the cross spectrum of such a pair of variables is readily determined, and the calculation shows that no simple interpretation of delay can be associated in general with this model. Another possible

⁸ Jenkins and Watts (ch. 8) and Granger and Hatanaka (ch. 5) contain a more detailed discussion of cross spectra.

TABLE 1—PHASE ANGLE AND τ

Implied by Exponentially Declining Weights Distributed Lag ($\beta = .7$)			
$2\pi f$ (radians)	Period ($1/f$)	Phase $-\alpha$ (radians)	$-\tau$ (time units of the sampling period)
.1	62.8	.226	2.26
.5	12.5	.716	1.43
.9	7.0	.771	.86
1.7	3.7	.567	.33
2.5	2.5	.262	.10

linear model arises when there is feedback between the variables. This model and the explicit cross spectral statistics that are determined by it have been discussed by C. W. J. Granger and by the present author. Here again there are generally no simple conclusions that can be derived about the time domain relationships of the variables solely on the basis of the phase shift.

An intuitive suggestion has been made by Granger and M. Hatanaka that in some cases one might regard low frequencies as reflecting "long-run" relationships, while high frequencies correspond to the "short-run." Suppose that the spectral decomposition of the time-series statistically isolates essentially independent factors governing long, intermediate, and short runs. Finally assume that for each "run" the relationship between the economic variables is a pure delay. Under these conditions, the computation of τ for the frequencies corresponding to different runs yield the pure delay for that run. The procedure adopted by Cargill seems to rest implicitly on some argument similar to the one just sketched. But this model seems very implausible for the wage and price variables on which his study is based.

While such a model is conceivable, there is a large burden of proof in establishing its relevance for a particular pair of economic series. The assumption that each frequency band with high coherence corresponds to an independent pure delay surely requires an explicit justification. One can imagine time-series containing strong periodicities from

seasonal factors, where the seasonal factors may be largely independent from other factors. Even here, some analysis is required to explain why a pure delay between the variables at the seasonal frequency is more plausible than a distributed lag.

The main conclusion of this analysis is that phase leads and lags measured from cross spectral estimates will rarely provide economists with direct estimates of the time domain relationships that are of interest. To avoid terminological confusion, it might be useful for economists using spectral techniques to reserve the unmodified words "lead" and "lag" exclusively for descriptions of phase relations, and to describe translation in the time domain by the phrase pure delay. If it is assumed in some application that phase shift or τ corresponds to pure delay, evidence should be provided to justify the assumption. If phase information is not intended to convey the impression of pure delay, an explicit warning is desirable. Finally, the discussion of all cross spectral analyses would be greatly improved if an explicit model (or class of models) is presented of the assumed dynamic relationship linking the variables.

APPENDIX

Calculations of τ for a Simple Input-Output Distributed Lag System

Suppose that input and output are related by the discrete time distributed lag of exponentially declining weights that often appears in econometric models: $w_k = \beta^k$ ($k=0, 1, \dots$). In this case

$$(3) \quad W(f) = \sum_{k=0}^{\infty} \beta^k e^{-i2\pi f k} = 1/(1 - \beta e^{-i2\pi f})$$

From this complex function one obtains the following formula for the phase:

$$(3') \quad \alpha(f) = -\arctan[(\beta \sin 2\pi f)/(1 - \beta \cos 2\pi f)]$$

Dividing this expression by $2\pi f$ gives the time lag, τ , in Cargill's terminology.

Table 1 shows that for the simple exponential lag, $\alpha(f)$ is always negative. As frequency increases, the phase angle initially increases in magnitude until $\cos 2\pi f = \beta$, and thereafter declines. τ itself is a declining function of f .

These results demonstrate that it is highly misleading to interpret τ as the lag between economic series linked by this input-output model. The time response of the system to changes in the input is given by the exponentially declining weights, and not by the phase angle or τ .

REFERENCES

- V. Bonomo and C. Schotta, "A Spectral Analysis of Post-Accord Federal Open Market Operations," *Amer. Econ. Rev.*, Mar. 1969, 59, 50-61.
- T. F. Cargill, "An Empirical Investigation of the Wage-Lag Hypothesis," *Amer. Econ. Rev.*, Dec. 1969, 59, 806-16.
- C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, July 1969, 37, 424-38.
- and M. Hatanaka, *Spectral Analysis of Economic Time Series*, Princeton 1964.
- J. C. Hause, "Leads, Lags and Spectral Analysis," *Econometrica*, Oct. 1964, 32, 687.
- G. M. Jenkins and D. G. Watts, *Spectral Analysis and its Applications*, San Francisco 1968.
- W. Kaplan, *Operational Methods for Linear Systems*, Reading, Mass. 1962.
- E. Malinvaud, *Statistical Methods of Econometrics*, Chicago 1966.
- T. J. Sargent, "Interest Rates in the Nineteen-Fifties," *Rev. Econ. Statist.*, May 1968, 50, 164-72.

Subsidized Housing in a Competitive Market: Comment

By GORDON TULLOCK*

Edgar Olsen's recent article in this *Review* is a significant contribution to clarifying the economics of a complex and difficult area. It is not the purpose of this comment to raise any questions as to his economic analysis, but to point out that, granted competition, it might be extremely difficult to subsidize low-income family consumption of superior housing by the method he suggests. In a sense, my objection is against interest, since I myself would much prefer that any subsidies on housing for lower income families use the Olsen method rather than the method of direct government provision. As I shall suggest later, however, there is another possible procedure which I regard as superior to either the provision of public housing or the Olsen subsidy.

For simplicity, assume that some poor person receiving subsidies under the Olsen procedure would normally spend \$60 a month on rent. He is permitted to purchase for \$60 a \$100 rent certificate which can then be used to rent superior housing. This amounts to giving him an income supplement of \$40, but attempting to compel him to use it for one particular purpose. His utility would be higher if you simply gave him the \$40 and permitted him to spend it on anything he chose. It seems reasonable if you did so he would indeed improve his housing, but would also improve his consumption of other goods as well. To use a rough rule of thumb, let us assume that if he were given \$40 in cash every month, he would choose to spend \$10 of this in increasing his consumption of housing—renting an apartment at \$70—and spend the other \$30 on other matters. Clearly, from the stand-

point of the poor person the receipt of a direct subsidy would be superior.

Granted that this is so and that the market is highly competitive (even if not perfectly competitive), it seems likely that the individual would be able to find a landlord who is willing to rent him an apartment which is normally worth \$70 for the \$100 certificate, and then make an under-the-table rebate to him of \$30. Olsen says that, "It would be illegal to exchange these certificates for other than housing services," but it seems to me that this is a type of crime which is extremely hard to detect. The only people involved would be the landlord and the poor tenant, and both would benefit from the crime. The so-called "crimes without victims," such as gambling, prostitution, drug sale, all present problems in enforcing the law. In fact many people, including myself, feel that these laws should be repealed. Olsen, in effect, is creating a new "crime without victim," and we can assume similar problems in enforcement. Further, intensive police activity to limit rebates might create a risk with the result that both the poor person and the landlord would be worse off than if the police activity did not exist, but most rebating would continue.

From the standpoint of the recipient, such a rent certificate is inferior to a direct cash payment. Surely there is some cash payment of slightly less cost to the state than the rent certificate which would be, from the standpoint of the recipient, superior to the certificate. Granting this, it seems to me that we should aim at direct cash payment. I presume that Olsen would agree. The reason he is advocating this particular mechanism is because he believes that subsidies aimed at increasing poor persons' consumption of housing services, rather than simply increas-

* Center for Study of Public Choice, Virginia Polytechnic Institute.

ing their income, are a more or less permanent part of our economy, and he wishes to make them more efficient. I cannot quarrel with this desire on his part, but I doubt that his particular technique would work. It seems to me that we would be better advised to try to change government policy toward raising the incomes of the poor rather than trying

to adjust their consumption toward the qualitative standards of those who are not poor.

REFERENCE

- E. A. Olsen, "A Competitive Theory of the Housing Market," *Amer. Econ. Rev.*, Sept. 1969, 59, 612-22.

Subsidized Housing in a Competitive Market: Reply

By EDGAR O. OLSEN*

Gordon Tullock has called to our attention the difficulty in preventing the recipient of a rent certificate from converting his certificate into the equivalent of a cash grant. His comments apply not only to certificates for housing but also to certificates for other goods, e.g., food stamps. Even though phrased in terms of housing, my reply has the same generality.

In the framework of Paretian welfare economics, the cost of enforcing the requirement that rent certificates be spent only on housing is relevant to determining the best means for redistributing consumption. Hence, Tullock *may* be correct in saying that this cost is so large that all transfers should be in the form of cash. However, this is not necessarily the case, as I demonstrate in the second part of my reply.

It is necessary to take exception to a belief explicit in Tullock's comment. He clearly argues that transfers in kind would result in inefficient resource allocation even if it were costless to prevent recipients from converting their in-kind subsidy into cash. He attempts to prove his point by showing that the recipient and his landlord could benefit by violating the provision of the certificate, and that no third party would be hurt by this violation.¹ He suggests that I advocate rent certificates because I believe that housing subsidies, though undesirable,

are inevitable and because I want the housing service consumed by subsidized families to be produced efficiently, as it would be by private producers in a competitive market. In fact, I think that housing subsidies and certain other in-kind transfers are desirable provided that their recipients can be costlessly prevented from converting the in-kind transfer into a cash grant.

In the first part of this reply, I will show that there exists a set of indifference maps and a societal budget constraint such that a costlessly enforced rent certificate scheme will result in efficient resource allocation. Allowing the recipient to violate the restriction on the use of his certificate will make someone worse off and will result in inefficient resource allocation. Existence is proven by means of an example in which there is a consumption externality. In this example, an infinite number of Pareto optimal allocations may be attained by rent certificate plans but not by unrestricted cash grants. One of these optima is the Lindahl solution associated with the pre-transfer distribution of income.² It is because I believe that consumption externalities are pervasive and because I attach special normative significance to the Lindahl solution that I proposed rent certificates. However, the case for rent certificates is stronger than this at least in my example because all Pareto optimal allocations at which all people are better off than they would be in the absence of transfers may be attained by rent certificates but not by unrestricted cash grants.

The assumptions and results of this example will be stated in the paper; a guide to

* Assistant professor of economics, University of Virginia. I am especially grateful to Stanley M. Besen, Edgar K. Browning, Joseph S. DeSalvo, Harold M. Hochman, Roland N. McKean, James D. Rodgers, and Gordon Tullock for comments which were helpful in revising an earlier draft of this paper.

¹ A similar argument attempting to show that it is better to give recipients unrestricted cash grants than to reduce the price that they pay for one good is due to Alan T. Peacock and D. Berry. Their argument about negative income and excise taxes follows easily from an earlier discussion of positive income and excise taxes which is reproduced and criticized by Milton Friedman. My objection to Tullock's proof is different from Friedman's.

² As Paul Samuelson (1969, p. 102) makes clear, a good generating a consumption externality is a public good by his definition. For example, if *A* is concerned about *B*'s housing, then housing consumed by *B* is a public good. See Richard Musgrave (pp. 73-78) for the Lindahl solution to the determination of the optimal quantity of a public good and the optimal distribution of taxes to pay for it.

the mathematical derivation of these results may be obtained from the author.

I. Optimality with Costless Enforcement

Assume that there are two individuals in society, the grantor and the recipient. They consume two goods, nonhousing X and housing H . The grantor directly consumes only nonhousing, but he also cares about the recipient's consumption of housing. The recipient directly consumes both nonhousing and housing. Let X_g and X_r be the quantities of nonhousing directly consumed by the grantor and recipient, respectively, and H_r be the quantity of housing directly consumed by the recipient. Suppose that the indifference maps of the grantor and the recipient are

$$(1) \quad X_g^a H_r^1 = a$$

$$(2) \quad X_r^b H_r^2 = b$$

where X_g , X_r , and H_r are nonnegative. Assume that nonhousing and housing are produced at constant costs of \$2 and \$1 per unit, respectively. Finally, assume that the grantor has an income of \$400 and the recipient an income of \$100 per time period.³ Hence, the budget frontier of this society is

$$(3) \quad 2(X_g + X_r) + H_r = 500$$

In this example there are infinitely many Pareto optimal consumption patterns within society's budget constraint. Samuelson (1954, 1955) would choose among these patterns by means of a social welfare function depending only on the utility indices of the members of society. This is not the only means of selecting one of the Pareto optimal alloca-

tions to be the grand optimum. Indeed, my belief in the desirability of rent certificates is based partly on a normative theory in the tradition of Lindahl in which the grand optimum is determined by another means. This theory is based upon individual preferences, the pretransfer distribution of income, and certain pricing rules.⁴

The normative pricing rules of the theory are as follows. If the good is a private good, then the theory says that each consumer should pay the marginal cost of producing it. In my example, nonhousing is a private good. Hence, both the grantor and the recipient should pay \$2 for each unit of nonhousing that they consume. If a good is a public good, then each person should pay a price for it such that the marginal rate of substitution between the public good and the private good is equal to the ratio of his price of the public good to the marginal cost of producing the private good. In my example, the mathematical representations of these rules are

$$(4) \quad \frac{\partial(X_g^a H_r^1)/\partial H_r}{\partial(X_g^a H_r^1)/\partial X_g} = \frac{P_g^h}{2}$$

$$(5) \quad \frac{\partial(X_r^b H_r^2)/\partial H_r}{\partial(X_r^b H_r^2)/\partial X_r} = \frac{P_r^h}{2}$$

where P_g^h is the price to be paid by the grantor and P_r^h is the price to be paid by the recipient for each unit of housing that the recipient consumes. These normative prices are the prices which determine the optimal taxes according to the benefit approach of Lindahl and Bowen. Equations (4) and (5) are equivalent to

$$(6) \quad 2X_g - 9P_g^h H_r = 0$$

$$(7) \quad X_r - 2P_r^h H_r = 0$$

Of course, the sum of the prices paid by different people for each public good must be equal to the marginal cost of producing it if, as assumed, the marginal cost is constant.

⁴ See Olsen (1969) for a recent elaboration of this theory.

³ To keep the exposition simple, I work with an exchange model and abstract from production. We might think of this society as composed of two retired people living on annuities and buying both goods in the perfectly competitive markets of another society. The production side could be considered in a trivial way by assuming that there is one factor of production, that one unit of nonhousing can be produced by two units of the input and one unit of housing by one unit of the input, and that the grantor owns 400 units and the recipient 100 units of the productive factor. In this case, equation (3) would be a production possibility frontier.

In my example,

$$(8) \quad P_o^h + P_r^h = 1$$

Finally, this normative theory assumes that each individual should consume only as much of the goods as can be bought with his initial income at the optimal prices. Therefore,

$$(9) \quad 2X_o + P_o^h H_r = 400$$

$$(10) \quad 2X_r + P_r^h H_r = 100$$

Equations (6) through (10) are the mathematical representation of the normative theory in this particular example. Therefore, we have five equations, only one of which is linear, and five unknowns, X_o , X_r , H_r , P_o^h , and P_r^h . Despite the nonlinearities there is only one solution ($X_o=180$, $X_r=40$, $H_r=60$, $P_o^h=2/3$, $P_r^h=1/3$) to this system of equations. Hence my grand optimum allocation of resources is ($X_o=180$, $X_r=40$, $H_r=60$). The utility indices of the grantor and the recipient at this allocation are about 161.3 and 43.4. It can be shown that this allocation of resources is one of the infinite number of Pareto optimal allocations given the indifference maps (1) and (2) and the societal budget constraint (3). It can also be proven that this allocation cannot be reached by cash grants alone. Therefore, subsidies in kind do not necessarily result in inefficiency.

In the absence of transfers between the grantor and the recipient, the grantor would spend all of his income on nonhousing. Hence, he would consume 200 units of this good. The recipient would choose to consume 40 units of nonhousing and 20 units of housing. The utility indices of the grantor and the recipient at this allocation are about 158.9 and 34.8. It can be shown that this allocation of resources is not Pareto optimal. Thus, there are circumstances in which transfers are necessary for efficient resource allocation.⁵

In this situation I would propose that the

⁵ This has already been proven in more general cases by Otto Davis and Andrew Whinston and by Harold Hochman and James Rodgers.

government sell to the recipient a rent certificate with a face value of \$60 and charge the recipient \$20. At this point in the analysis, I assume that the government can costlessly force the recipient to use his certificate for housing only, and I propose that the government do so. Though the recipient is free to spend more than \$60 on housing, he will not choose to do so in this case. Hence, my proposals would result in the recipient's consuming 60 units of housing and 40 [= (\$100-\$20)/\$2] units of nonhousing. I propose that the government collect \$40 in taxes from the grantor. The grantor would buy 180 units of nonhousing with the \$360 that he would have left. With \$20 from the recipient and \$40 from the grantor, the government redeems the certificate from the seller of housing at face value. These proposals would result in my grand optimal allocation of resources. This example proves that there are situations in which rent certificates will result in efficient resource allocation.⁶

Suppose that we allowed the recipient to redeem his rent certificate at face value for cash but prohibited further transfers. The grantor would continue to consume 180 units of nonhousing. The recipient will exchange his certificate for \$60 and will have \$140 [= \$100-\$20+\$60] in cash to spend as he pleases. In this case, he would consume 56 units of nonhousing and 28 units of housing. The utility indices of the grantor and the recipient at this allocation are about 149.4 and 48.8. Naturally, the recipient prefers this allocation of resources to my grand optimum. However, the grantor prefers my grand optimum. Indeed, in this particular example he prefers the allocation in the absence of transfers to this allocation. Since it is this type of model which leads me to recommend rent certificates, I cannot agree with Tullock that my restriction on the use

⁶ This is not to say that rent certificates are necessary for efficient resource allocation but only that they are sufficient. My grand optimum could be attained by lowering the price per unit of housing to the recipient to one third of a dollar and allowing him to consume any quantity that he chooses. Furthermore, there is one Pareto optimal allocation in this example that can be attained by an unrestricted cash transfer.

of the rent certificate creates a crime without a victim. The grantor is the victim. Furthermore, the allocation ($X_g=180$, $X_r=56$, $H_r=28$) is not Pareto optimal. Both the recipient and the grantor would be better off, for example, at the feasible consumption pattern ($X_g=170$, $X_r=440/9$, $H_r=560/9$). Therefore, to allow recipients of subsidies in kind to convert these subsidies into cash may result in inefficient resource allocation.

In this example there is only one Pareto optimum ($X_g=0$, $X_r=200$, $H_r=100$) that can be attained by a cash transfer between the grantor and the recipient.⁷ Therefore, even in the presence of consumption externalities, subsidies in kind are not necessary for efficient resource allocation. Since the grantor would consume nothing directly in this situation, he would be as bad off as he could be (i.e., his utility index would be zero). Under these circumstances, I doubt that anyone other than the recipient would argue for unrestricted cash grants.

Attainment of the infinity of other Pareto optimal allocations requires noncash grants. An infinite subset of this infinite set of Pareto optima is composed of allocations which both the recipient and the grantor prefer to the allocation in the absence of transfers. Each of these allocations can be attained by a voluntary rent certificate scheme. In a forthcoming paper, I have proved that if the recipient prefers one of the consumption patterns which the rent certificate scheme makes available to him to his pretransfer pattern, then either the rent certificate is equivalent to an unrestricted cash grant or the recipient will spend precisely the face value of his certificate on housing. Since the infinite subset consists only of allocations not attainable by cash transfers, the second possibility obtains. Therefore, if the allocation (X_g^* , X_r^* , H_r^*) is a Pareto optimal allocation preferred by both the grantor and the recipient to the

pretransfer allocation of resources, then this allocation can be attained by offering to sell to the recipient a rent certificate with a face value of $\$H_r^*$ for $\$100 - \$2X_r^*$ and taxing the grantor $\$H_r^* - \$100 + \$2X_r^*$.

II. Optimality with Costly Enforcement

Throughout the previous section, I assumed that the recipient could be forced costlessly to spend the face value of his rent certificate on housing. I now dispense with this assumption.

We know that the recipient would prefer to exchange his certificate for its face value in cash. Therefore, unless some attempt is made to enforce the provision on the use of his rent certificate, he will violate it. In order to deter the recipient from violating this provision, there must be some positive probability of being caught. To create this probability some resources must be expended on law enforcement. If resources are expended on law enforcement, then the allocation ($X_g=180$, $X_r=40$, $H_r=60$) is not attainable because some of the money spent on nonhousing and housing must be diverted to law enforcement.

In my example the violation of the restriction on the use of the rent certificate is a crime with a victim. When should this crime be prevented?⁸ On grounds of efficient resource allocation, a crime should be prevented if, and only if, the loss to the victim exceeds the gain to the criminal by more than the minimum cost of preventing the crime. In my example, the loss to the grantor due to the recipient's violation would be \$29.24. That is, the grantor is willing to pay up to this amount to prevent the recipient from violating the restriction by exchanging his certificate for cash at face value. The gain to the recipient from this violation would be \$15.43. That is, if the restriction were not enforced, then the recipient would be as well off as he would have been had he been given an unrestricted cash grant of \$40; if the restriction were enforced, then the recipient would be as well off as he would have

⁷ In the many goods, many persons case, there will be one such point for each person who is concerned only about goods that he consumes directly because the allocation of resources that would result from transferring all of society's income to such a person is Pareto optimal.

⁸ Gary Becker and Harold Demsetz have investigated this general question in some detail.

been with a cash grant of \$24.57. Therefore, if the cost of enforcing this provision is more than \$13.81 [= \$29.24 - \$15.43], then on grounds of efficient resource allocation the provision should not be enforced. In this case I would agree with Tullock that the recipient should be given a cash grant. However, if the cost of enforcement is less than \$13.81, then the provision should be enforced. In this case, I would say that an enforced rent certificate plan is preferable to a cash grant.

III. Concluding Comments

Having shown that rent certificates can be justified in the framework of Paretian welfare economics and that there is some cost of enforcing the restriction on the use of the certificates so low that rent certificates are preferable to cash grants, I now wish to agree with Tullock that the cost of enforcing this provision is likely to be very high. As a result, I think that it is entirely reasonable to argue for transfers in cash even though consumption externalities are pervasive. This, however, is an empirical question. It cannot be settled solely on theoretical grounds.

I proposed rent certificates because I believe that there are many paternalistic altruists in this country and that housing is one of the goods that these people think the poor value too lightly.⁹ My belief stems from the casual observation that most governmental and nongovernmental transfers to the poor are in kind (e.g., public housing, food stamps, and medicare). If this sort of consumption externality proves to be unimportant, then I will withdraw my rent certificate proposal.

⁹ A discussion of other consumption externalities that have been postulated to justify housing subsidies appears in my dissertation (pp. 19-21, 38-48).

REFERENCES

- G. S. Becker, "Crime and Punishment: An Economic Approach," *J. Polit. Econ.*, Mar./Apr. 1968, 76, 169-217.
- O. A. Davis and A. B. Whinston, "Welfare Economics and the Theory of the Second Best," *Rev. Econ. Stud.*, Jan. 1965, 32, 1-14.
- H. Demsetz, "The Exchange and Enforcement of Property Rights," *J. Law Econ.*, Oct. 1964, 7, 11-26.
- M. Friedman, "The 'Welfare' Effects of an Income Tax and an Excise Tax," *J. Polit. Econ.*, Feb. 1952, 60, 25-33; reprinted in M. Friedman, *Essays in Positive Economics*, Chicago 1953, 100-113.
- H. M. Hochman and J. D. Rodgers, "Pareto Optimal Redistribution," *Amer. Econ. Rev.*, Sept. 1969, 59, 542-57.
- R. A. Musgrave, *The Theory of Public Finance*, New York 1959.
- E. O. Olsen, "A Welfare Economic Evaluation of Public Housing," unpublished doctoral dissertation, Rice Univ. 1968.
- , "A Normative Theory of Transfers," *Publ. Choice*, spring 1969, 6, 39-58.
- , "Some Theorems in the Theory of Efficient Transfers," *J. Polit. Econ.*, forthcoming.
- A. T. Peacock and D. Berry, "A Note on the Theory of Income Redistribution," *Economica*, Feb. 1951, 18, 83-90.
- P. A. Samuelson, "The Pure Theory of Public Expenditure," *Rev. Econ. Statist.*, Nov. 1954, 36, 387-89.
- , "Diagrammatic Exposition of a Theory of Public Expenditure," *Rev. Econ. Statist.*, Nov. 1955, 37, 350-56.
- , "Pure Theory of Public Expenditure and Taxation," in J. Margolis and H. Guitton, eds., *Public Economics*, New York 1969.

Expectations and the Demand for Bonds: Comment

By RICHARD ROLL*

John H. Wood concluded his recent article in this *Review* by stating

... the expectations hypothesis [of the term structure of interest rates] is logically invalid ... the Hicks and Lutz equations rely on sub-optimal decision rules and, consequently, are without behavioral significance. ... The awkward Hicks and Lutz formulations have hindered inquiries into the effects of uncertainty on the term structure ... our results have made it possible for the study of the structure of rates under such conditions to proceed on a sound theoretical basis. [pp. 529-30]

The purposes of this comment are to determine the historical accuracy of Wood's interpretation of Irving Fisher (pp. 273-74), J. R. Hicks (pp. 144-47), and F. A. Lutz (pp. 499-529) and to examine the economic validity of Wood's theory.

I. Fisher, Hicks, and Lutz

According to Wood, the traditional expectations hypothesis of the term structure implies a decision rule of the following form: A trader¹ should "... prefer n -period bonds, be indifferent between n - and one-period bonds, or prefer one-period bonds, when"

$$(1) \quad (1 + R_n) \begin{matrix} \geq \\ \leq \end{matrix} \cdot [(1 + R_1)(1 + {}_1r_1) \dots (1 + {}_{n-1}r_{n-1})]^{1/n} \quad (\text{p. 524.})$$

where R_k is the current observed market yield on k -period bonds and ${}_jr_1$ is the yield currently expected by the trader to prevail on one-period bonds j periods hence.

At first glance, expression (1) might seem

* Carnegie-Mellon University

¹ We shall assume (with Wood) that this trader operates in a bond market free of transaction costs, maximizes return over a horizon as long or longer than the longest-term bond outstanding, and is indifferent to risk.

an accurate algebraic representation of the expectations hypothesis. A one-period forward rate applicable j periods hence is, by definition,

$$(1 + {}_{j-1}r_1) \equiv \frac{(1 + R_j)^j}{(1 + R_{j-1})^{j-1}}$$

so that a current market yield in terms of forward rates is

$$(1 + R_n) \equiv [(1 + R_1)(1 + {}_1r_1) \dots (1 + {}_{n-1}r_{n-1})]^{1/n}$$

The "pure" expectations hypothesis states that each observed market forward rate is equal to its corresponding expected future spot rate, (see David Meiselman, p. 10) ${}_jp_k = {}_jr_k$, so that the equality form of (1) is the market equilibrium equation.² Indeed, Wood indicates this fact and then destroys our belief in the logic of the expectations hypothesis by presenting an alternative decision rule which earns a higher expected return.

The trader using Wood's decision rule will "... prefer n -period bonds, be indifferent between n - and m -period bonds, or prefer m -period bonds, when"

$$(2) \quad (1 + R_n) \begin{matrix} \geq \\ \leq \end{matrix} (1 + R_m)^{m/n} = \frac{(1 + {}_1r_{n-1})^{(n-1)/n}}{(1 + {}_1r_{m-1})^{(m-1)/m}} \quad (\text{p. 525.})$$

Decision rule (2) is based on a strategy of maximizing return one period at a time. In the first period, period 0, the investor should make pairwise comparisons of all bonds and select the bond whose expected holding period yield during the first period is highest. Wood proves that decision rule (2) is superior to (1). Rule (1) is sub-optimal because it prompts the trader, who is indifferent to risk, into basing current decisions partly on an-

² This interpretation of (1) requires the r 's to be rates expected by some hypothetical composite market trader.

anticipated future decisions. There is no doubt that (2) is a rule superior to (1). The question, however, is whether decision rule (1) can be ascribed to the expectations hypothesis.

An examination of Fisher, Hicks, and Lutz will show that no such decision rule was ever recommended by them. In fact, Fisher and Lutz presented the same example that Wood used to show the inferiority of rule (1). This example, to be discussed in detail in the next section, involves a trader who expects a rise in the rate on long-term bonds. As Wood shows, such a trader will prefer short-term bonds now (p. 527) and will wait until rates have risen to buy long-term bonds. Compare Fisher's analysis:

Those who expect the [long-term] rate of interest to fall will prefer to invest in long-time securities at the present market rates, even when those rates are less than on securities of shorter time, *while those who expect the [long-term] rate of interest to rise will prefer short-time securities.* (italics added) [p. 274]

and Lutz':

The second possibility is that the investor may expect the yield on the bond at some intermediate date to exceed the average of the short rates from that date onwards, i.e., he expects the market price of the bond to be relatively *low* at that date. He will then contemplate going into the short market now and into the long market later. [p. 514-15]

Fisher and Lutz were clearly aware that decision rule (1) leads to incorrect actions.

Hicks, the third author Wood associates with rule (1), should probably be left out of the discussion entirely. His theory of the term structure is intricately connected to risk which we have assumed away here.³ Let it suffice to note that decision rule (1) cannot be found in his book.

Finally, we must mention that all three

writers, Fisher, Hicks, and Lutz, were discussing the term structure as a market equilibrium phenomenon and not as a normative theory for the guidance of investors. It is quite easy to be misled by the equality form of (1), which is the market equilibrium condition, into accepting the inequalities, which have nothing whatever to do with the theory. A thorough reading of these early articles will demonstrate the unfairness of Wood's criticism, "Discussions of the expectations theory of the term structure of interest rates have tended to be rather mechanical, ignoring the microeconomic foundations of market equilibrium solutions" (p. 522). Nothing could be less true. All three authors owe the frequent references to their work to a *concentration* on the foundations of market equilibrium and a style of expression that is lucid and *non-mechanical*.

II. A Revised Bond Investor's Decision Rule

We now turn from Wood's interpretation of history to a discussion of his decision rule. This section intends to prove that neither Rule (1) *nor* Rule (2) is optimal.

The correct decision rule can be demonstrated with Wood's three-period example. He assumed that two risk-indifferent traders, *G* and *H*, held expectations depicted by⁴

$$(H.1) \quad (1+R_2)^2 = (1+R_1)(1+r_1)$$

$$(H.4) \quad (1+R_3)^3 > (1+R_1)(1+r_1)(1+r_2)$$

$$(G.4) \quad (1+R_3)^3 < (1+R_1)(1+r_2)^2$$

The lower case *r*'s denote the trader's expected future spot rates and upper case *R*'s denote market spot rates at period 0. Trader *G*'s inequality (G.4) was Wood's decision rule (2) and Trader *H*'s inequality (H.4) was intended to be the decision rule (1) implied by the expectations hypothesis. Wood assumed, via (H.1), that both *G* and *H* were indifferent between one- and two-period bonds.⁵ According to these decision rules, *H* prefers three-period over one-period bonds and *G* prefers the opposite. Wood showed

³ Lutz was also concerned with risk and devoted much of his article to its discussion. Hicks, however, discussed practically nothing else.

⁴ The *H* and *G* notations are Wood's.

⁵ Both *H* and *G* were assumed to have made a decision to commit their funds for at least three periods.

that G 's one-period gain would indeed be greater than H 's. H is led into error because he "... expects at time 0 to prefer two-period over one-period securities at time 1. Because of his use of decision rule (2), he is influenced by this expected future preference in his portfolio decision at time 0" (p. 527).

Neither the rule Wood recommended (2) nor the rule he attributed to the expectations hypothesis (1) will lead to optimal actions by a risk-indifferent trader. The correct rule is (a) *make one-period spot loans now with available resources* and (b) *make forward loans now if the forward rate is greater than the corresponding expected future spot rate*. In the present example, the forward rates are

$$(3) \quad (1 + {}_1\rho_1) = (1 + R_2)^2 / (1 + R_1)$$

$$(4) \quad (1 + {}_2\rho_1) = (1 + R_3)^3 / (1 + R_2)^2$$

$$(5) \quad (1 + {}_1\rho_2)^2 = (1 + R_3)^3 / (1 + R_1)$$

where ${}_k\rho_k$ is the k -period forward rate to begin j periods hence.

To determine the optimum investment at time zero, the trader must compare forward rates to expected spot rates as follows: ${}_1\rho_1$ to ${}_1r_1$, ${}_1\rho_2$ to ${}_1r_2$, and ${}_2\rho_1$ to ${}_2r_1$.

Using equations (H.1) and (3), we obtain for the first comparison,

$$(6) \quad 1 + {}_1r_1 = 1 + {}_1\rho_1$$

which indicates that now, in period 0, the trader is indifferent between making one-period forward loans to begin one period hence and waiting until period 1 to make one-period spot loans.

Using the inequality (G.4) and equation (5), we have

$$(1 + {}_1\rho_2)^2 = \frac{(1 + R_3)^3}{(1 + R_1)} < (1 + {}_1r_2)^2$$

or

$$(7) \quad {}_1\rho_2 < {}_1r_2$$

Since the two-period forward rate to begin one period hence is less than the two-period expected spot rate, the trader should issue forward loans now. He should borrow forward. In a world of perfect capital markets and zero transaction costs, he can do this by selling three-period bonds short and buying

one-period bonds with the proceeds. At the beginning of period 1, he will receive an expected capital gain of

$$(8) \quad d_1 \left[(1 + R_1) - \frac{(1 + R_3)^3}{(1 + {}_1r_2)^2} \right]$$

where d_1 is the dollar amount of three-period bonds sold short and one-period bonds purchased. By referring to inequality (G.4), one can verify that the expected gain, represented by (8), is indeed a positive quantity.

Using equations (H.1) and (4) and inequality (H.4), we obtain the third comparison,

$$(1 + {}_2\rho_1) = \frac{(1 + R_3)^3}{(1 + R_2)^2} > 1 + {}_2r_1$$

or

$$(9) \quad {}_2\rho_1 > {}_2r_1$$

Inequality (9) implies that the trader should now make one-period forward loans for two periods hence. Again, in the perfect world of this example, the transaction can be accomplished by selling short two-period bonds and using the proceeds to buy three-period bonds. *After two periods*, this will bring an expected capital gain of

$$(10) \quad d_2 \left[\frac{(1 + R_3)^3}{(1 + {}_2r_1)} - (1 + R_2)^2 \right]$$

where d_2 is the dollar amount of both the long and short transaction. By substituting for $1 + R_2$ from (H.1) and using inequality (H.4), one can verify that (10) is also positive.

In summary, the present decision rule instructs a trader to make the following transactions at period zero:

- (a) Buy one-period bonds with available resources.
- (b) Sell short three-period bonds and use the proceeds to buy one-period bonds.
- (c) Sell short two-period bonds and use the proceeds to buy three-period bonds.

Transaction (b) is kept open for one period, period 0 to 1, and transaction (c) is kept open

for two periods. Only transaction (a) was recommended by Wood's decision rule (2) and none of the three transactions were recommended by (1). The positive expected capital gains of (8) and (10) that accrue to a trader using the revised decision rule prove that rules (1) and (2) are sub-optimal.

The quantities of bonds bought and sold in transactions (b) and (c) are unspecified. This is a very important fact that requires elaboration. If the trader is truly risk indifferent, and really wants to maximize *expected* return, he would attempt to make d_1 in expression (8) and d_2 in (10) as large as possible. Only by transacting an infinite quantity of bonds would he maximize expected return. The fact that we rarely observe investors attempting to trade infinite amounts brings out the unrealism of the preceding example. Markets are not perfect in the special sense used there and traders can neither grant nor issue unlimited quantities of loans. Even if they could, it is likely that none *would* because no trader operating with unlimited liability is completely indifferent to the risk of total ruin.

The example is important, however, in clarifying the central point of this comment: Expected future spot rates do have behavioral significance. By comparing them to current forward rates, the bond trader chooses an optimal investment strategy.

With the introduction of uncertainty, their comparison to forward rates acquires an even more crucial role. Assuming no risk of default, the forward rate is perfectly certain whereas the corresponding future spot rate is a random variable.⁶

REFERENCES

- I. Fisher, *The Nature of Capital and Income*, London 1912.
- J. R. Hicks, *Value and capital*, 2d ed., Oxford 1946.
- F. A. Lutz, "The Structure of Interest Rates," *Quart. J. Econ.*, Nov. 1940, 40, 36-63, rep. in AEA, *Readings in the Theory of Income Distribution*, Homewood 1951, 499-529.
- D. Meiselman, *The Term Structure of Interest Rates*, Englewood Cliffs 1962.
- R. Roll, "The Efficient Market Model Applied to U. S. Treasury Bill Rates," unpublished doctoral dissertation, Univ. Chicago 1968, forthcoming as *The Behavior of Interest Rates: An Application of the Efficient Market Model to U.S. Treasury Bill Rates*, New York 1970.
- J. Wood, "Expectations and the Demand for Bonds," *Amer. Econ. Rev.*, Sept. 1969, 59, 522-30.

⁶ The implications of this environment have been worked out in my doctoral thesis.

Expectations and the Demand for Bonds: Comment

By A. BUSE*

The traditional theory of the term structure of interest rates, as exemplified by J. R. Hicks, is usually considered to be a coherent and consistent doctrine. It is not the logic of the theory that has been questioned by the critics but its empirical relevance; expectations to Kingdom Come being the more humorous manifestation of this attitude. However, in a recent stimulating paper in this *Review*,¹ John Wood has asserted that the decision rule implied by the traditional theory leads to sub-optimal results (as measured by holding period returns) unless the investor is assumed to hold each security until maturity. Without this assumption, Wood argues, the optimal decision rule requires that investors make forecasts of interest rates for only one period into the future, unlike the traditional theory which involves forecasts of expected one-period rates for n periods into the future. A three-period example is used to demonstrate the result.

The purpose of this comment is to demonstrate that Wood's sub-optimality result does not hold. Although the algebra of his three-period example is correct, the result he obtains depends on the unstated assumption that the Wood-type investor makes correct forecasts and the traditionalist does not. This being the case, it is hardly surprising that the traditionalist is sub-optimal. If the forecasting abilities are reversed, so are the optimality relations.

Wood's three-period example, (pp. 526-28), can be used to demonstrate the assertions made in the previous paragraph. The point of departure is a pair of inequalities (H.4) and (G.4) which specify the divergence between existing market rates and those of

the H -investor (traditional decision rule) and the G -investor (Wood decision rule).

$$(H.4) (1 + {}_0R_3) > [(1 + {}_0R_1)(1 + {}_1r_1)(1 + {}_2r_1)]^{1/3}$$

$$(G.4) (1 + {}_0R_3) < (1 + {}_0R_1)^{1/3}(1 + {}_1r_2)^{2/3}$$

The current one- and three-period market rates² are ${}_0R_1$ and ${}_0R_3$. As of period zero, ${}_1r_1$ and ${}_2r_1$ are the one-period rates investor H expects in periods one and two. It follows that the H -investor expects the two-period rate in period one to be

$$[(1 + {}_1r_1)(1 + {}_2r_1)]^{1/2} - 1$$

The G -investor, who makes forecasts of rates for only one period into the future, expects the two-period rate to be ${}_1r_2$ in period one. G and H are assumed to have identical expectations about the one-period rate in the next period.

Given these conditions, investor H will buy three-period bonds and G will buy one-period bonds. Wood then compares the returns obtained by H and G after the lapse of one period. Obviously, the return to G is ${}_0R_1$ per unit of investment. The return to H will depend on the price of the three-period security at the start of period one. If this price is denoted by ${}_1P_3$, the one-period return to H is given by

$$({}_0P_3 / {}_1P_3) - 1 = [(1 + {}_0R_3)^3 / (1 + {}_1R_2)^2] - 1$$

Wood (p. 527) states that the return to H is given by

$$(1 + {}_0R_3)^3 / (1 + {}_1r_2)^2 - 1,$$

which is less than ${}_0R_1$ by (G.4). Wood's result follows if ${}_1r_2 = {}_1R_2$; that is, G 's expectation of the two-period rate for period one is the rate that actually prevails in the market

* Associate professor of economics, University of Alberta.

¹ Although I try to demonstrate that Wood's results are in error, I found the paper stimulating nonetheless because I was forced to think carefully about the implications of the traditional expectations theory.

² A minor change in notation from the original has been made in order to make the argument as explicit as possible. Thus, ${}_tR_j$ is the observed yield on a j -period bond in period t . Similarly, ${}_tP_j$ stands for the observed price of a j -period bond in period t .

at $t=1$. In short, G 's expectations are correct. If this is the case, then H 's expectation of the two-period rate must be wrong since by (H.4) and (G.4). (Wood's inequality (7))

$$(1) \quad {}_1r_2 > [(1 + {}_1r_1)(1 + {}_2r_1)]^{1/2} - 1$$

where the right-hand side is H 's expectation of the two-period rate.

Assume now that H 's expectations are correct so that after a lapse of one period $(1 + {}_1R_2)^2 = (1 + {}_1r_1)(1 + {}_2r_1)$. Then the holding period return to H is given by

$$(2) \quad [(1 + {}_0R_3)^3 / (1 + {}_1r_1)(1 + {}_2r_1)] - 1$$

which by (H.4) is greater than ${}_0R_1$. The G -investor now has the lower rate of return and his decision rule appears to be sub-optimal.

Another possible way of recognizing Wood's implicit assumption is to consider his description (p. 527) of the inequality (7)

$$(3) \quad (1 + {}_1r_2)^2 > (1 + {}_1r_1)(1 + {}_2r_1)$$

which is implied by (H.4) and (G.4). He states that "Inequality (7) shows where H 's expectations differ from the forward rates implicit in current rates." Clearly he is treating ${}_1r_2$ as the implied market forward rate (to be subsequently realized) when in point of fact the current rates imply a two-period rate in period one equal to³

$$[(1 + {}_0R_3)^3 / (1 + {}_0R_1)]^{1/2} - 1$$

³ It is of course possible that the "market" has correct expectations so that $(1 + {}_0R_3)^3 / (1 + {}_0R_1) = (1 + {}_1R_2)^2$. In this case the realized returns to H and G are equal although H 's realized return is less than his expected return. G could be similarly disappointed if the inequalities in (H.4) and (G.4) were reversed.

which by (G.4) is not equal to ${}_1r_2$. It can also be noted that inequality (7) is inconsistent with Wood's assumption that H and G have identical expectations but this inconsistency is secondary to the main point that the G decision rule does not possess the dominance property attributed to it by Wood.

To sum up, Wood has not proved that a decision rule based on forecasts of the long rates in the next period is superior to a decision rule which forecasts the expected one-period rates. It follows that he has not demonstrated that the traditional expectations theory is incomplete without the assumption of bond holding to maturity. These conclusions are not that surprising since Wood indicates that the decision rules lead to different actions only when inequalities such as (7) hold. Such inequalities specify differences in forecasts of future rates and the results of any action will depend on the success of the forecast. The higher return will always accrue to the better forecaster. This is not to say, however, that the decision rule proposed by Wood does not have substantial intuitive appeal. Such a rule may in fact be relevant to a successful explanation of the behavior of the term structure. But this is an empirical issue, not a logical one. Whether the issue will be resolved next year or by Kingdom Come is another matter.

REFERENCES

- J. R. Hicks, *Value and Capital*, 2d ed., Oxford 1946.
 J. H. Wood, "Expectations and the Demand for Bonds," *Amer. Econ. Rev.*, Sept. 1969, 59, 522-30.

Expectations and the Demand for Bonds: Comment

By REUBEN A. KESSEL*

The thesis of John Wood's article is: Predicting the next period price of a bond constitutes a more correct criterion, i.e., leads to better results for the investor than predicting the expected spot rates over the life of a bond which is the classical view. It is the purpose of this note to argue that this constitutes a distinction without a difference and the dissimilar results Wood obtains in comparing the two criteria is a consequence of his failure to understand the process of arbitrage and speculation in the world he postulated, a world of zero transactions costs.

Wood's proof that his criterion is better than the criterion he imputes to Fisher, Lutz, and Hicks rests on the following example: If the average of the expected one-period rates one and two periods hence is less than the expected two-period rate one period hence, then directly estimating the price of a bond one period hence will yield a different and better guide to action than estimating the expected one-period rates into the indefinite future.

Symbolically,¹

$$(1) \quad (1 + {}_1r_2)^2 > (1 + {}_1r_1)(1 + {}_2r_1)$$

In addition,

$$(2) \quad (1 + R_3) > [(1 + R_1)(1 + {}_1r_1)(1 + {}_2r_1)]^{1/3}$$

and

$$(3) \quad (1 + R_3) < (1 + R_1)^{1/3}(1 + {}_1r_2)^{2/3}$$

According to Wood, Hicks et al. will choose the three-period rate and earn

* Professor of business economics, Graduate School of Business, University of Chicago.

¹ The expected rate is r . The prescript represents the number of periods that must elapse before a rate becomes a spot rate; the postscript represents the number of periods a rate encompasses. R depicts spot rates.

$\frac{(1 + R_3)^3}{(1 + {}_1r)^2}$ in the first period, whereas Wood

would buy one-period bonds and earn $(1 + R_1)$

which, from (3), is greater than $\frac{(1 + R_3)^3}{(1 + {}_1r_2)^2}$.

It is at this point that Wood commits a crucial error. In a market in which there are zero transactions costs, the inequality between the geometric mean of expected one-period rates one and two periods hence and the expected two-period rate one period hence cannot hold; this market is not in equilibrium. Under the conditions postulated, arbitrage will occur. Arbitrageurs will buy forward two-period money for delivery one period hence and sell forward for delivery one and two periods hence one-period money. As a result, the inequality will be eliminated. This arbitrage opportunity holds in a world of uncertainty as well as perfect forecasting. Once the inequality is eliminated, both criteria will yield identical answers.²

It should come as no surprise that both criteria yield identical answers. The processes by which investment decisions are made are identical. To obtain the expected price one period hence, one must look at the forward rates over the life of the securities outstanding. This is the same process that Lutz, Hicks, and Fisher postulate.³

² This same point is relevant for understanding the implications of the other inequalities Wood considered. Forward rates undervalued in the market will be bought and conversely.

³ There are important differences between Hicks and Lutz that Wood ignores. Uncertainty is integral to the views of Hicks; liquidity preference and certainty cannot logically coexist.

Expectations and the Demand for Bonds: Reply

By JOHN H. WOOD*

The first half of Richard Roll's comment contains an argument that my interpretation of the traditional expectations hypothesis of the term structure of interest rates is incorrect. The second half argues that both the decision rule that I infer from the traditional theory and the decision rule that I have advanced as an improvement upon that approach are sub-optimal. Adolf Buse and Reuben A. Kessel, on the other hand, accept my interpretation of the received theory and argue that the traditional approach is correct. I will respond first to Roll's criticism of my interpretation (and that of Buse and Kessel) of the traditional literature.

I. Fisher, Hicks, and Lutz

Any dispute concerning Irving Fisher's understanding of the decision process by which investors choose among bonds of different maturities ought to be settled by quoting his complete statement, of which Roll gives us the final sentence. The passage immediately preceding the sentence quoted by Roll is as follows:

If the intention in advance is to reinvest, it becomes important not simply to know the present rate of interest, but to forecast the future rate. This enters into the calculations of an investor who holds a 25-year bond at 5 percent. He will usually regard the final payment as 'principal', intending that when it becomes due it shall be reinvested in a similar 25-year bond. He, therefore, is

not really buying a 25-year income stream of \$5 a year plus \$100 at the end of the term, but is buying let us say, a 50-year income stream consisting of \$5 per year for the first 25 years and an *unknown* amount per year during the second 25 years. In order to forecast what income will be received in the second period, he has to forecast the rate of interest. In other words, although the bond represents nominally a fixed and certain series of income items, yet, in view of the intention to reinvest, it actually represents an income which is quite uncertain after 25 years, because of the uncertainty in the future rate of interest. Such an investor, if he expected the rate of interest at the end of 25 years to be 2 percent, would, in purchasing the above-mentioned bond, be getting \$5 a year for 25 years and \$2 a year for the next 25 years. Under these conditions, if he could buy a 50-year bond at 4 percent, he would prefer to do so. But, if he expected the rate of interest to remain, for each 25-year period, at 5 percent, he would prefer, rather than invest now in a 50-year bond at 4 percent, to invest in the 25-year bond at 5 percent, intending to reinvest at 5 percent at the expiration of the term. His forecast of what the rate of interest will be in 25 years will thus materially affect the choice of investments to-day.¹ [1906, pp. 273-74]

Fisher's discussion clearly indicates that in choosing between current purchases of long- and short-term bonds the investor forecasts interest rates many years in the future. He then compares the current rate

* Finance department, University of Pennsylvania. I wish to take the opportunity in this reply to remedy an omission from my article. Herschel Grossman has published a paper in which he also develops at some length the point that, under conditions of certainty and zero transactions costs, an investor who desires to maximize wealth over any horizon utilizes forecasts of interest rates only one period in the future. Unfortunately, Grossman's contribution came to my attention too late for me to refer to it in my article.

¹ This quotation also makes clear that, when Fisher writes about expectations of rises and falls in the rate of interest, he is referring to the rate of interest on 25-year bonds that is expected to prevail 25 years in the future. In Fisher's example, this is the short-term rate. Hence, Roll is incorrect when he inserts "long-term" in brackets to modify "rate of interest" in the first sentence of his quotation.

on long-term bonds with the average of current and future rates on short-term bonds. If the former exceeds the latter, he buys long-term bonds; if the relationship is reversed, he buys short-term bonds; if the equality holds, he is indifferent between long- and short-term bonds.² Such an action based upon such a comparison is the essence of a decision rule. So far as I am aware, this passage from Fisher is by far the clearest and most complete description of the decision process underlying the equilibrium relationship among interest rates implied by the expectations hypothesis. The quality of discussions of that theory, even by Fisher,³ deteriorates rapidly after 1906.⁴ For example, after stating his assumptions of certainty, no transactions costs, and complete shiftability among bonds by both borrowers and lenders, Lutz gives the equilibrium relationship among rates that he asserts to be implied by these assumptions in a footnote with no intervening discussion (pp. 499–500). The equation given by Lutz (who credits Fisher, Hicks, and others with the development of the theory) is simply an algebraic expression of Fisher's argument for the special case in which the equality holds. The remainder of Lutz's Section I, which is devoted to the basic theory (i.e., where certainty and zero transactions costs prevail), consists solely of a series of algebraic manipulations of the formula contained in his footnote with the added implicit assumption of zero elasticity of expectations. Roll's quotation from the Lutz article, which is taken from Section IV (pp. 512–20) is not

² Compare Fisher's discussion with my statement of the traditional theory (pp. 522–24) which was drawn directly and without modification from the passage just quoted and from the equivalent (when the equality holds) passage in J. R. Hicks (pp. 144–45).

³ See (1930, p. 70) in which Fisher merely asserts without defense and without reference that "a rate on a five year contract may be considered as a sort of an average of five theoretically existing rates, one for each of the five years covered."

⁴ The most notable exception is Burton Malkiel (pp. 18–20), who discusses the decision process implied by the traditional theory in detail. J. W. Conard (p. 294) also alludes to a decision process of the type attributed in my article (pp. 523–24) to that theory.

relevant to the criticism presented in my paper, which is directed toward the Fisherian theory that Lutz accepted without question in his Section I (pp. 499–504). The discussion in Section IV of Lutz presents an analysis of the determination of the equilibrium structure of rates in a world in which Fisher's theory has been modified by the introduction of transactions costs (in Section II) and uncertainty (in Section III).⁵ Consequently, the passage quoted by Roll is taken from the discussion of a theory very different from that of Fisher's and very different from that which has been accepted as "the" expectations hypothesis, the cornerstones of which are the assumptions of certainty and zero transactions costs. It will further become clear to the reader of Lutz's discussion in its entirety that one of the similarities between Sections I and IV is that the investor is required to forecast short rates several periods in the future, though not necessarily in the analysis of Section IV over the entire life of the long-term investment. Contrary to the impression given by Roll's discussion and his quotations from Fisher and Lutz, both of these writers require investors to forecast rates several periods into the future.

Roll's assertion with respect to Hicks' treatment of the term structure of interest rates is also incorrect. As the reader may verify, Hicks (pp. 144–45) begins with a statement of the equilibrium relationship among rates in the absence of risk. This relationship is identical to that implied by Fisher's discussion ("if no interest is to be paid until the conclusion of the whole transaction" (Hicks, p. 145)) and is expressed in its general form in equation (1) of my article. It is no accident that the terminology preceding my equation (1) is very similar to that used by Hicks. Only after his statement without qualification of the theory received

⁵ The precise manner in which uncertainty and transactions costs enter Lutz's modification of the traditional theory is not, however, at all clear. Lutz is extremely vague about the assumptions underlying the analysis of his Section IV. See D. G. Lockett for a valiant attempt to decipher this portion of Lutz's article.

from Fisher does Hicks (like Lutz) introduce uncertainty into the analysis.⁶

II. Roll's Decision Rule

Roll refers to the three-bond example of my article, where I showed that an investor using the decision rule of the traditional theory "prefers three-period over one-period securities," whereas an investor following the decision rule that I advance, which utilizes forecasts only one period in the future, "prefers one-period over three-period securities" (p. 527). Roll agrees with my point that the investor should "(a) make one-period spot loans now with available resources." But he goes further than I did in his recommendation (b) that the wealth maximizing investor should also issue three-period loans now in order to buy one-period bonds with the proceeds. I agree. This is a useful extension of my results. Actions (a) and (b) are both desirable if Roll's expression (8) is positive, which follows from the inequality (G.4) on page 527 of my article. (Also see the inequality on page 528.) As I emphasized, (G.4) utilizes forecasts only one period in the future.

Roll's recommendation (c), however, is sub-optimal because at this point he commits the classic Fisher-Hicks-Lutz error. He uses the Hicksian decision rule (H.4) to derive the result that the investor should issue two-period bonds and use the proceeds to buy three-period bonds. "*After two periods*, this

⁶ Hicks' attempt to deal with the term structure under uncertainty, like that of Lutz, was unsuccessful. Hicks imposed "liquidity premiums" on a theory which has since been shown to be invalid (see Wood) and with no discussion of the objective functions to be maximized by investors or of the nature of the constraints confronting them. With one exception, those studies of the term structure under uncertainty that explicitly use techniques of constrained optimization have been limited in effect to comparisons between one- and two-period bonds (see G. O. Bierwag and M. A. Grove, and H. A. J. Green), which is the only case in which the traditional theory is valid (see Wood p. 526) and, hence, the only case which can support the kinds of rigorous analysis under uncertainty such as have been pursued by Bierwag and Grove and by Green. The exception is Grossman who is able to deal with bonds of varying maturity because he recognizes that the traditional theory is invalid and discards it from the beginning of his analysis.

will bring an expected capital gain" equal to the amount shown in (10). But Roll has already shown in (a) and (b) that one-period bonds are currently preferable to three-period bonds. Why should an investor issue three-period bonds in order to raise funds with which to buy one-period bonds, as in (b), and then in (c) issue two-period bonds in order to buy three-period bonds?⁷ Action (c) is consistent with wealth maximization only if the investor is bound to hold securities for two-periods. Otherwise, the investor will never buy three-period bonds currently regardless of the source of funds. Roll seems to be on the verge of an explicit recognition of this condition when he says that action (c) brings an expected capital gain "*after two periods*." But if we constrain investors to hold securities until maturity so that (c) is optimal, then (a) and (b) cease to be optimal. This is precisely the point made in my article.⁸

Roll's point in his next-to-last paragraph is valid and is one that I recognized in my article.⁹ The infinite elasticity of demand functions that follows from the assumptions of certainty, zero transactions costs, and perfect competition is one of the more troublesome aspects both of the traditional theory and of the theory advanced in my article. In the case of divergent expectations, market equilibrium solutions are rendered indeterminate unless "We assume . . . that the command over resources by any individual investor is limited . . ." (p. 523, fn. 5). On the other hand, as Meiselman (pp. 10, 54, 57) has argued, perfectly elastic demand functions may actually strengthen the theory if expectations are identical, which is an assumption usually associated with the expectations hypothesis (see Lutz, p. 499, Conard, p. 290, Malkiel, p. 18, and Wood, p.

⁷ Note that in (a) and (b), Roll prefers one-period to three-period bonds and then in (c) prefers three-period to two-period bonds when it is clear from (H.1) that by both the Hicksian and one-period forecasting rules, the investor is indifferent between one- and two-period bonds. He gets these non-transitive results because he jumps back and forth between contradictory decision rules.

⁸ See the last paragraph of Section II, p. 528.

⁹ See fn. 5, p. 523. Also see Don Patinkin p. 68.

522). But Roll's suggestion that our understanding of the structure of interest rates will be improved by the introduction of uncertainty into our analysis is good advice. On this point, see footnote 6 above and the last paragraph of my article.

III. Buse: Contradictory Expectations?

Buse argues that in my three-bond example investors pursue different courses of action not because of different decision rules but because of differences in forecasting. It is difficult to respond to this criticism except by pointing out that I explicitly assumed identical expectations (p. 527). Further, it is easily seen that the substitution of some configuration of expected rates into the decision rule (2) that Buse agrees is implied by the traditional theory (p. 524) may yield a different result from that obtained by substituting the same expectations into the decision rule (5) that utilizes forecasts only one period in the future (p. 525). My three-bond example is merely "the simplest case in which decision rules (2) and (5) produce different results . . ." (p. 526). Can anyone, Buse excepted, doubt that different decision rules may yield different results even though the same information is available to both?

Buse obtains his results that investor *G* sometimes does better and sometimes worse than *H* by indiscriminately jumping between the assumption that (1) is an inequality and the assumption that (1) is an equation. Confusion results especially from his substitution of the equation form of (1) into the inequality (H.4) and then comparing the result with the inequality (G.4). Buse performs this comparison despite my demonstration (p. 526) that, given identical expectations and indifference between one- and two-period bonds by *H* and *G* (assumptions explicitly underlying both my equilibrium and disequilibrium examples), if (1) is an equation then (H.4) and (G.4) reduce to equations (H.2) and (G.2) which are equivalent. Naturally, Buse's different and sometimes contradictory assumptions yield contradictory results. I retained the same set of assumptions throughout the course of my discussion of the disequilibrium case (in-

cluding the assumption that the inequality (1) holds) and hence obtained a unique result. Buse's mistake is the common one in discussions of the expectations hypothesis of confusing equilibrium and disequilibrium situations.

IV. Kessel: Identical Result in Equilibrium

Buse's last line of defense consists of the same argument as that advanced by Kessel. That argument is that the decision rule implicit in the traditional theory and the one suggested in my article yield identical results in equilibrium, i.e., when the inequalities in my example become equations. Kessel's argument consists of the following four statements: (i) The example discussed in my paper describes a position of market disequilibrium; (ii) Forces will be set in motion that will restore equilibrium, in which case (1)-(2) become equations; (iii) In equilibrium, both the Hicksian and one-period forecasting rules "yield identical answers"; (iv) Consequently, "to obtain the expected price one period hence, one must look at the forward rates over the life of the securities outstanding." Kessel and I are in complete agreement with respect to the first three statements.¹⁰ But the fourth statement is a non sequitur. One wonders why, even in equilibrium, if the two criteria yield identical answers it is necessary to forecast rates many periods into the future when by Kessel's admission the one-period forecasting model does as well. In fact, any decision criterion, including a dart board, yields an optimal result in equilibrium.

Neither Kessel nor Buse perceives the nature of the distinction from the standpoint of decision makers between equilibrium and

¹⁰ It was shown in my article that, during the period of adjustment, an investor using the one-period forecasting rule will earn at least as much and sometimes more than an investor using the traditional rule, a result not disputed by Kessel. For discussions of movements from disequilibrium to equilibrium positions that are consistent with Kessel's description of the adjustment process, see pages 523-24, 528-29 of my article. Also see Malkiel pp. 19-20. Kessel's statement (iii), which has served for many years as the clinching argument in support of the traditional theory, is discussed in my article on pages 524, 526-27.

disequilibrium situations. Under the assumptions stated at the beginning of my article, there is no investment problem when the market is in equilibrium (i.e., when inequalities such as (1)–(2) are eliminated) and hence no need for an investment criterion because all investments are expected to yield identical one-period returns. In a world in which all markets were somehow continuously in equilibrium, neither the Hicksian nor the one-period forecasting rule would be of any use; neither would have any behavioral significance and the forward rates implicit in both would be mere tautologies. Consequently, the two rules are meaningful only in regimes in which disequilibria are possible. And it was shown in my article that, under the assumptions generally thought to underlie the traditional theory, the one-period forecasting rule dominates the rule implicit in the traditional theory in disequilibrium situations. Thus, under conditions of certainty and zero transactions costs, wealth maximizing investors make decisions on the basis of forecasts only one period in the future. The result is that the forward rates more than one period in the future defined by the observed structure of rates are without behavioral significance.

An analogy drawn from the theory of the firm will illustrate my point regarding the importance of distinguishing between market equilibria and disequilibria in the development of decision rules. In long-run competitive equilibrium where entry is free, profit maximizing firms will produce quantities such that marginal and average costs are equal to each other and to the price of the product. If the industry were continuously in long-run equilibrium, the following decision rules would yield identical and optimal results: (a) operate such that marginal cost equals price (subject to the condition that marginal cost is rising); (b) minimize average cost. But to say that (a) and (b) yield identical results in equilibrium does not obscure

the fact that in disequilibrium situations (i.e., nearly all of the time), (a) dominates (b).¹¹

REFERENCES

- G. O. Bierwag and M. A. Grove, "A Model of the Term Structure of Interest Rates," *Rev. Econ. Statist.*, Feb. 1967, 49, 50–62.
- J. W. Conard, *An Introduction to the Theory of Interest*, Berkeley 1959.
- I. Fisher, *The Nature of Capital and Income*, New York 1906.
- , *The Theory of Interest*, New York 1930.
- H. A. J. Green, "Uncertainty and the Expectations Hypothesis," *Rev. Econ. Stud.*, Oct. 1967, 34, 387–98.
- H. I. Grossman, "Expectations, Transactions Costs, and Asset Demands," *J. Finance*, June 1969, 491–506.
- J. R. Hicks, *Value and Capital*, 2d ed., Oxford 1946.
- R. A. Kessel, "The Cyclical Behavior of the Term Structure of Interest Rates," Nat. Bur. Econ. Res. Occas. Paper 91, New York 1965.
- D. G. Lockett, "Professor Lutz and the Structure of Interest Rates," *Quart. J. Econ.*, Feb. 1959, 73, 131–44.
- F. A. Lutz, "The Structure of Interest Rates," *Quart. J. Econ.*, Nov. 1940, 55, 36–63; rep. in *AEA Readings in the Theory of Income Distribution*, Homewood 1946, 499–529.
- B. G. Malkiel, *The Term Structure of Interest Rates*, Princeton 1966.
- D. Meiselman, *The Term Structure of Interest Rates*, Englewood Cliffs 1962.
- D. Patinkin, *Money, Interest and Prices*, 2d ed., New York 1965.
- J. H. Wood, "Expectations and the Demand for Bonds," *Amer. Econ. Rev.*, Sept. 1969, 59, 522–30.

¹¹ With respect to Kessel's fn. 3, see my response above to Roll's criticism of my reference to Hicks.

Output of the Restrained Firm: Comment

By A. ROSS SHEPHERD*

In a recent article in this *Review*, Milton Z. Kafoglis examined the price and output behavior of the restrained monopoly firm and found that when firms seek to maximize output or scale of operations, "... the output of the restrained monopoly may exceed that predicted by existing models and may even be pushed beyond optimum (in the Paretian sense) as a result of sales at prices below marginal cost" (p. 583). *Inter alia*, Kafoglis finds that even in single markets under increasing cost Pareto optimal output "... will be exceeded by the output-maximizing firm and if demand elasticity exceeds unity in the range of restraint, by the revenue-maximizing firm" (p. 586). The purposes of this comment are: (1) to reveal the implicit assumption on which Kafoglis' result is based; (2) to show for the increasing cost case the assumption under which constrained output and revenue maximizers will produce the optimal output; and (3) to show that even in single markets *monopsony* power will likely result in sub-optimally *large* outputs for constrained output and revenue maximizers.

Figure 1 shows the monopolist's average revenue (AR) and marginal revenue (MR) curves together with various long-run cost curves emanating from point M .¹ Compared to Kafoglis' Figure 1 (p. 584), curve MA corresponds to the rising portion of Kafoglis' average cost curve (AC) and MB gives marginal cost (MC) above average cost. Thus in

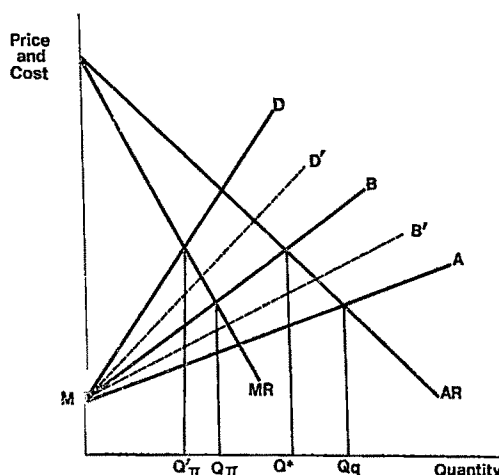


FIGURE 1

Figure 1, MA is the long-run average cost curve and MB is the curve marginal to MA . According to Kafoglis, Q_π is the profit-maximizing output, Q^* is the Pareto optimal output, and Q_q is the maximum output consistent with a break-even profit constraint. Now Q^* is Pareto optimal only if MB gives the cost to *society* of producing and selling the marginal unit. In order for Q_π to be the profit-maximizing output, MB must also give the increment to the *firm's* total cost that results from the production and sale of the Q th unit. These two conditions are consistent if, and only if, the firm pays no rent to factors of production. This is possible if factor prices are invariant with respect to the level of the firm's activity; it is also possible if the firm is a perfectly discriminating monopsony. From the general context of Kafoglis' discussion it would appear that he implicitly assumes fixed input prices.

It seems clear that an adequate analysis of industry behavior must include the possibility that industry expansion occurs under conditions of unconstrained substitution among inputs and rising factor supply

* Associate professor of economics, University of Missouri-Kansas City. I wish to thank the managing editor for helpful comments.

¹ As Dean A. Worcester, Jr. has pointed out (fn. 9, pp. 879-80), the convention of drawing U-shaped cost curves for the monopoly firm can be misleading. The monopoly firm is the analytical counterpart of the competitive industry, and the long-run cost curves in each case will likely come from the same family. For some purposes the conventional U-shaped curve is convenient in that it illustrates all the long-run cost possibilities. However, the present discussion is limited to the increasing cost case, so we show only rising cost curves.

prices. When rising average costs reflect rising input prices the monopolist is altogether likely to recognize that he has monopsony power, which means that he will see that his marginal costs include increases in intramarginal factor rents. From society's standpoint, however, these intramarginal rent increases are not costs; only the cost requisite to induce the marginal factors to produce and market the marginal unit of output is a cost to society. The implications of rising factor supply prices for the behavior of the restrained monopoly firm are clearly seen after an appropriate reinterpretation of Figure 1.

In the context of rising factor prices, curve *MA* is average cost exclusive of factor rents; *MB*, which is mathematically marginal to *MA*, is both marginal cost exclusive of factor rents, and average cost including average competitive factor rents. Under perfect competition *MB* is the long-run average cost or supply curve of the industry, and *MA* is analytically irrelevant because no competitive firm can avoid paying factor rents. In the presence of monopsony power, however, *MA* is to be interpreted as the average cost curve of the perfectly discriminating monopsonist, for whom *MB* is the curve of the marginal costs. Curve *MD* is marginal to *MB* and gives under perfect competition the total increase in cost, including the increase in intramarginal factor rents, associated with the production and sale of the *Q*th unit.²

Consider now the familiar, illustrative

² It is worthwhile to specify the assumptions that underlie the interpretations of this paragraph. 1) Factors are homogeneous as viewed by the industry for whom the cost curves are drawn, but heterogeneous from the standpoint of alternative uses. This ensures rising marginal transfer prices and the payment of factor rents to intramarginal units. 2) Expansion of the industry does not increase the prices of the factors in alternative uses so that the transfer cost of any given unit of a factor is independent of the level of the industry's output. 3) The industry production function exhibits constant technical returns to scale, and there are no pecuniary economies of large-scale production. These latter two assumptions ensure coincidence between the curves showing marginal competitive cost exclusive of factor rents and average competitive cost including average factor rents. The *loci classici* for all this are Joan Robinson's chapters 8 and 10.

case of a product produced with two factors. It is well known that a least-cost input combination requires:

$$(1) \quad \frac{MPP_1}{MFC_1} = \frac{MPP_2}{MFC_2},$$

where *MPP* denotes marginal physical product, *MFC* is marginal factor cost and the subscripts indicate the respective factors. The relationship between *MFC*, factor price (*P*) and the factor supply elasticity (*E*) is also well known:

$$(2) \quad MFC_i = P_i(1 + 1/E_i) \quad (i=1, 2) (E \neq 0)$$

In the case of perfect competition, each firm's perspective is such that for it $E_i = \infty$ and $MFC_i = P_i$, and the invisible hand will lead the competitive industry to use inputs such that $MPP_1/P_1 = MPP_2/P_2$. Now if supply elasticities are not infinite but are equal, ratios of *MFC*'s still equal ratios of *P*'s and least-cost combinations for given outputs are not altered by perceived monopsony power. In this case the monopoly-monopsonist's cost curves will coincide with those of the competitive industry (provided that monopolization of a competitive industry does not yield other net private economies or diseconomies). In Figure 1 the profit-maximizing monopoly-monopsony will aim at output *Q* while the constrained output and revenue maximizers will produce the optimal output, *Q**. We have it, then, that when rising output costs reflect increasing input prices, the constrained output and revenue maximizers will produce a socially optimal rate of output if factor supply elasticities are equal. In what follows it will be shown that for this optimal result to obtain equal factor supply elasticities are necessary as well as sufficient.

Factor supply elasticities will usually differ and, as compared to the competitive industry, the monopoly firm with monopsony power (hereafter, simply "monopoly") will economize on the relatively inelastic factor. This will both reduce the monopolist's (private) costs and misallocate resources. For some rate of output *Q* let:

$$(3) \quad \frac{MPP_1}{P_1(1 + 1/E_1)} \neq \frac{MPP_2}{P_2(1 + 1/E_2)},$$

where

$$(4) \quad \frac{MPP_1}{P_1} = \frac{MPP_2}{P_2}$$

A competitive least-cost combination obtains, and the output is produced with the socially optimal combination of inputs because the competitive industry has ignored intramarginal factor rents, which are not social costs. However, the monopolist will note that he can reduce the private cost of this output by economizing on the less elastic factor until extra output per last dollar spent is the same for each input. Hence, if at every point on the competitive supply curve factor supply elasticities differ, the monopolist's average cost will be below the competitive level for all outputs greater than zero.

In Figure 1 MB' and MD' are, respectively, the monopolist's average and marginal cost curves when factor supply elasticities differ. It is clear that monopsonistic private cost saving will lead the constrained

output and revenue maximizers beyond Q^* , the social optimum, while the profit maximizer will move closer to Q^* .³ The ultimate in monopsonistic cost saving is realized by the perfectly discriminating monopsonist, for whom the marginal cost is MB and average cost is MA . This brings us back to the zero rent results of Kafoglis.⁴

REFERENCES

- M. Z. Kafoglis, "Output of the Restrained Firm," *Amer. Econ. Rev.*, Sept. 1969, 59, 583-89.
 J. Robinson, *The Economics of Imperfect Competition*, London 1933.
 D. A. Worcester, Jr., "Pecuniary and Technological Externality, Factor Rents, and Social Cost," *Amer. Econ. Rev.*, Dec. 1969, 59, 873-85.

³ In the case of the profit maximizer the output adjustment resulting from this monopsonistic misallocation of resources counteracts in some measure the social distortion of monopsonistically contrived scarcity.

⁴ It will be noted that if the firm is both a perfectly discriminating monopsony and a perfectly discriminating monopoly, the profit maximizer will produce a socially optimal rate of output.

Output of the Restrained Firm: Reply

By MILTON Z. KAFOGLIS*

A. Ross Shepherd correctly reasons that the monopoly firm examined in my original analysis must expand with fixed input prices and that the average cost curve must, therefore, exclude rent to fixed factors. In relaxing this assumption to allow for rising input price (and the payment of full rent), Shepherd introduces the possibility that the output (or revenue) maximizing firm may produce the Pareto optimal output. We should conclude therefore that in the case of single markets under increasing cost the output maximizing monopoly will attain Pareto optimal output if the average cost curve coincides with the competitive supply schedule but will exceed Pareto optimal output if the average cost curve lies below the competitive supply schedule. In the ordinary case we would predict a lower average cost curve because the firm probably can avoid the payment of full rent and may practice various forms of discrimination in factor markets.

It would seem that the possibility of rent avoidance is especially significant in the case of regulated industries where effective calcu-

lations of opportunity costs are frustrated by legal and regulatory barriers which prevent consideration of alternative uses and by rate base calculations in terms of historical money costs. In the case of publicly-owned utilities this situation may be aggravated by tax exemptions and other concessions. Supported by such institutions the firm may be expected to develop a cost schedule which fails to reflect the full opportunity costs of the resources employed and, if restrained only by the fair return criterion, will be able to expand output beyond Pareto optimal.

Shepherd's comment is especially instructive because it brings attention to the fact that the level of the average cost curve is affected by industrial structure. However, since most of these effects will be intra-marginal, the marginal cost curve may not be altered significantly and analysis on the traditional assumption of profit-maximization may yield correct price and output predictions. On the other hand, output and price may be affected significantly by the level of the average cost curve in the case of restrained firms. In such cases it becomes necessary to be more explicit about intra-marginal factor payments than my original treatment.

* Professor of economics, University of Florida.

Production Indeterminacy with Three Goods and Two Factors: A Comment on the Pattern of Trade

By DOUGLAS B. STEWART*

James Melvin's examination of the indeterminacy in the three-good, two-factor, two-country trade model prompts him to claim in his recent article in this *Review* that whenever all goods are traded, that country exporting the labor intensive good will also be exporting the capital intensive good (p. 1263). Recognizing the damage this claim does to the standard Heckscher-Ohlin theorem, Melvin reformulates the theorem into a much weaker proposition.

We will show that Melvin's claim does not hold in general; that it is true if, and only if, both countries have identical relative factor endowments—a definitely uninteresting case. The example from which Melvin generalizes is often a *possibility* when endowment ratios differ, and this possibility alone is sufficiently damaging to the Heckscher-Ohlin theorem to merit comment. But, as we shall see, the damage is much less than Melvin would have it.

I. Notation and Assumptions

We assume a world in which each country has the (same) technology to produce three goods, and each has a fixed endowment of two factors, labor and capital. For the production functions we write:

$$(1) \quad X_1 = F_1(K_1, L_1)$$

$$(2) \quad X_2 = F_2(K_2, L_2)$$

$$(3) \quad X_3 = F_3(K_3, L_3)$$

These are assumed to be homogeneous of degree one, and to exhibit positive but decreasing marginal productivity in each factor. Both factors are assumed to be fully employed. Thus,

* Assistant professor of economics, University of Dayton. This paper was written while the author was a National Science Foundation Graduate Fellow at the University of Oregon. He thanks Chulsoon Khang for helpful suggestions.

$$(4) \quad L = L_1 + L_2 + L_3$$

$$(5) \quad K = K_1 + K_2 + K_3$$

for each country. By k we denote a country's endowment capital-labor ratio, K/L .

II. The Trade Pattern With Two Countries

In the three-good, two-factor, two-country case, the indeterminacy in production and trade appears when the terms of trade and each country's endowment ratio are such that it is possible for each country to produce all three goods. Taking the isoquant approach, the situation can be depicted as in Figure 1. Here, the isoquant of unit value output (in terms of, say, good 1) for each good is tangent to the unit isocost line ab , and the endowment ratio rays, k^a and k^b , lie in the cone coe . Except for the endowment rays, the same diagram serves for both countries since the production functions are assumed to be identical.

Suppose in equilibrium country A produces good 2 on a scale indicated by point f , and country B produces good 2 on a scale indicated by point h . Then we can determine by parallelogram construction that country A will necessarily produce goods 1 and 3 on a scale indicated by points g and j , respectively, and country B will necessarily produce goods 1 and 3 on a scale indicated by points m and n , respectively.¹ Assuming both countries have similar tastes, it is clear country B will export goods 2 and 3 and import good 1.² Thus, the depicted situation is

¹ The parallelogram is the geometrical representation of vector addition. To satisfy the factor employment assumption in Figure 1 for, say, country A we must have $\overrightarrow{of} + \overrightarrow{og} + \overrightarrow{oj} = \overrightarrow{op}$, where the bar denotes a vector. Given \overrightarrow{op} and \overrightarrow{of} we find \overrightarrow{og} by construction such that $\overrightarrow{of} + \overrightarrow{og} = \overrightarrow{op}$. We then find \overrightarrow{oj} such that $\overrightarrow{og} + \overrightarrow{oj} = \overrightarrow{op}$.

² By "similar tastes" we mean demand conditions do not differ enough at any relevant terms of trade to cause the larger producer of any good to import it. Melvin

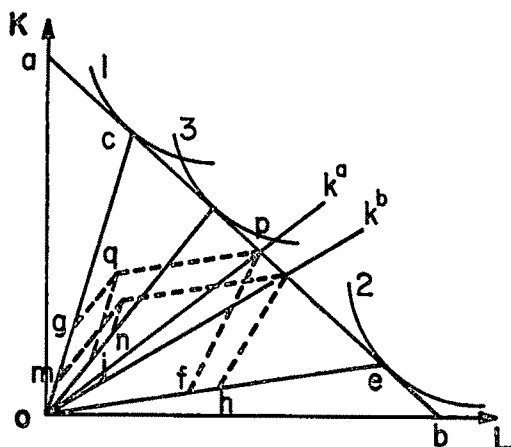


FIGURE 1

one in which country *B* is exporting the labor intensive good and importing the capital intensive good. We have, then, constructed a counterexample to Melvin's claim that the country which exports the labor intensive good will also export the capital intensive good. By manipulating the diagram one can convince oneself that the counterexample *can always* be constructed as long as $k^a \neq k^b$ and both are in the cone *coe*. Thus, $k^a = k^b$ is necessary for Melvin's proposition.

Suppose both countries have the same endowment ratio. Figure 2 illustrates this situation. Obviously, whichever country produces at point *a* and, hence, exports good 2, will also produce at point *c* and export good 1. Clearly, then, $k^a = k^b$ is a sufficient condition for Melvin's proposition. Therefore, we have shown that identical relative factor endowments is a necessary and sufficient condition for Melvin's proposition. A corrected version of his claim is: if both countries have the same endowment ratio, then country *X* exports the labor (capital) intensive good if and only if country *X* exports the capital (labor) intensive good.

We find no error in Melvin's construction of the production possibility surface. It appears he simply overlooked those possible

avoids explicit demand assumptions, but assumes (p. 1254) countries are identical except for their capital stocks.

trade patterns contrary to his claim but consistent with the standard Heckscher-Ohlin theorem, for our counterexample is easily illustrated with his own diagram (p. 1262) which we reproduce in Figure 3.

Here the goods are numbered such that good 2 is labor intensive and good 1 is capital intensive. Keeping Melvin's consumption point *T*, we have selected different production points *S* and *R*. We see the country producing at *R* produces goods 1 and 2 in quantities represented by points *b* and *d*, respectively, whereas the similar points for the country producing at *S* are *a* and *c*. Therefore, the country producing at *S* exports the labor intensive good and imports the capital intensive good, contrary to Melvin's claim.

If both countries have the same endowment ratio, *JH* and *J'H'* will coincide and contain the consumption point *T*, in which case both countries could produce at *T* and trade would be unnecessary. For this reason we feel this special case is of little importance.

III. The Heckscher-Ohlin Theorem

On the basis of his claim Melvin turns to a weak form of the Heckscher-Ohlin theorem. According to this version a country's export bundle will be intensive, relative to its import bundle, in that factor which the country holds in relative abundance. This statement is true whether endowment ratios are

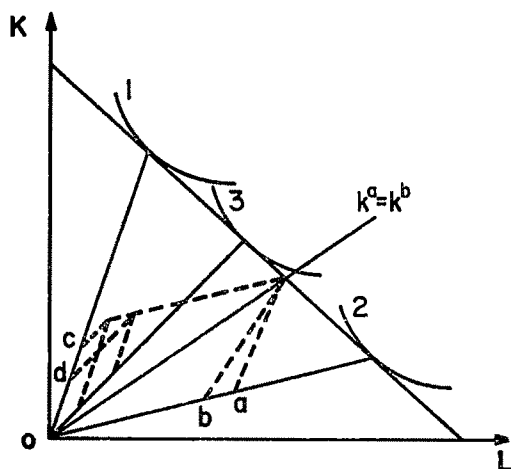


FIGURE 2

equal or not. If they are equal ($k^a = k^b$), Melvin's claim holds and his Heckscher-Ohlin theorem tells us the capital-labor ratio of all traded bundles is the same. If endowment ratios are unequal the theorem is still true, but because Melvin's claim is untrue the theorem is unnecessarily weak.

We will now develop a stronger version of the Heckscher-Ohlin theorem to apply to the indeterminacy case when endowment ratios are different. In this situation there are two possibilities: either 1) each country must produce either the capital intensive or labor intensive good, or alternatively 2) one country's endowment ratio is equal to the capital-labor ratio of the good of intermediate intensity and it produces only that good. This second possibility would be rather unusual and the pattern of trade is easily resolved, for obviously the other country will export both the capital intensive and labor intensive goods.

Figure 4 gives an example of the first possibility. Here both countries must produce good 2. If country A were to produce only goods 2 and 3, they would be produced on a scale represented by points a and n , respectively. Similarly, points b and j are the appropriate points if only goods 2 and 1 are produced. Thus the interval ab represents country A's range of possible outputs of

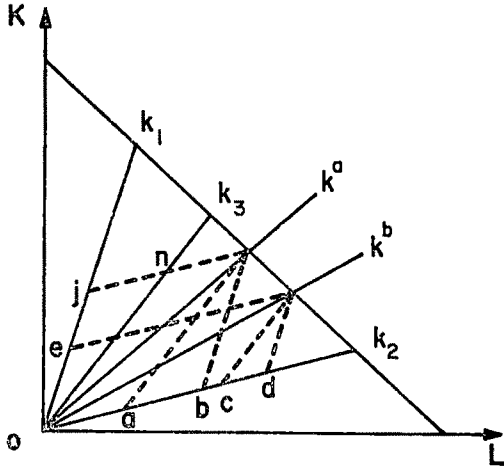


FIGURE 4

good 2. For country B this interval is cd . Since the intervals do not overlap country B must export good 2. It is always possible to draw k^a and k^b between k_3 and k_2 such that ab and cd do not overlap. For good 1, oe and oj represent the production ranges of countries B and A, respectively. Since $oe < oj$ country A is more likely to export good 1 than is country B. If k^a and k^b were closer together in the same price situation ab and cd would overlap and allow the possibility of country A exporting good 2. Also, the ratio oe/oj would be larger, thus decreasing the likelihood of country A exporting good 1.

Another example of the first possibility is the situation where k^a and k^b are on opposite sides of k_3 . In this case it turns out that if k^a and k^b are sufficiently different the production intervals of the two countries will not overlap for either good 1 or good 2, and the pattern of trade is that predicted by the standard Heckscher-Ohlin theorem. We leave the construction of this example to the reader.

Generalizing from examples of this type, we state our theorem for the indeterminacy case with $k^a \neq k^b$ and excluding possibility 2) above. As $|k^a - k^b|$ increases, the probability of each country exporting the good using intensively that country's abundant factor increases; if both countries must produce different goods, there is a value of $|k^a - k^b|$

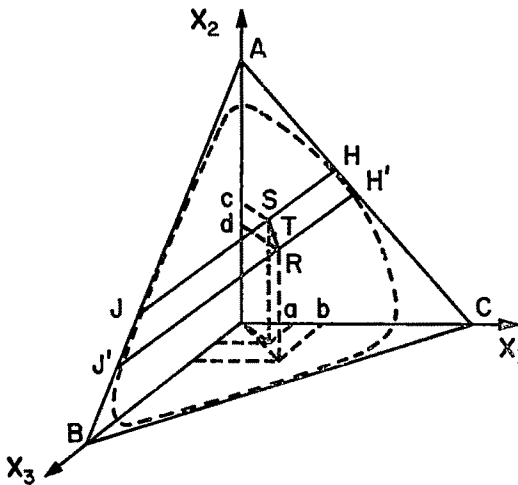


FIGURE 3

above which each country must export the good using intensively its abundant factor; if both countries must produce the *same* good, there is a value of $|k^a - k^b|$ above which that good must be exported by the country abundant in its intensively used factor.

Our theorem contains significantly more information than Melvin's, yet allows for the contradiction of the standard theorem which he points out. For sufficiently different

endowment ratios (remaining within the indeterminacy case) it gives the standard result that each country will export the good using intensively that country's abundant factor.

REFERENCE

- J. R. Melvin, "Production and Trade with Two Factors and Three Goods," *Amer. Econ. Rev.*, Dec. 1968, 58, 1249-68.

Production Indeterminacy with Three Goods and Two Factors: Reply

By JAMES R. MELVIN

Let me begin by agreeing in part with Douglas Stewart's argument. My claim that whenever all goods are traded, the country exporting the labor intensive good will also export the capital intensive good, does not hold in general. While the case I refer to is possible, I am, as Stewart has shown in his Section II, certainly guilty of a careless and incorrect generalization.

At the same time, Professor Stewart is guilty of much the same kind of error, for his statement that my claim "... is true if, and only if, both countries have identical relative factor endowments ..." (p. 241) is incorrect. Indeed one would be tempted to interpret this remark as a slip of the pen except for the fact that two similar statements are made subsequently. After correctly presenting a counterexample to my claim he says "Thus, $k^a = k^b$ is necessary for Melvin's proposition," (p. 242) and later "Therefore, we have shown that identical relative factor endowments is a necessary and sufficient condition for Melvin's proposition" (p. 242). Stewart himself presents a counterexample to all these statements when he says "The example from which Melvin generalizes is often a *possibility* when endowment ratios differ, ..." (his italics). Obviously if my claim is a possibility when endowment ratios differ, equal endowment ratios cannot be a necessary and sufficient condition for my claim. That the situation I have described is possible can be seen from Stewart's Figure 4. If endowments are such that ab and cd overlap, then it is possible for either country to export both the capital intensive and the labor intensive good while importing the intermediate one.

From my analysis I drew the conclusion that for the three-good, two-factor case, the traditional statement of the Heckscher-Ohlin Theorem must take the somewhat

weaker form that a country will export a bundle of goods which is intensive in the factor with which that country is relatively well endowed. In light of the fact that my example is still a possibility, I see no reason to change my conclusion. Stewart's argument has shown that the more traditional trade pattern is possible, but that is quite consistent with my reformulation of the theorem.

In his Section III, Stewart attempts to present a stronger version of the theorem than the one I presented. Unfortunately his analysis suffers from his failure to appreciate the fact that a possibility other than the two he mentions can exist; namely that a country may export both the capital intensive and the labor intensive goods while importing the intermediate one. Furthermore, while I must confess that I have had some difficulty understanding his theorem,¹ it does not seem to add much to our knowledge. If my interpretation of the spirit of his theorem is correct it states that the more different are the endowments of the two countries, the more likely is a country to export the commodity which is intensive in the factor with which that country is relatively well endowed. While this is certainly correct it is not particularly surprising, and unless precise conditions can be derived which will tell us when it is *not* possible for one country to export both the capital intensive and the labor intensive goods, we

¹ I do not understand what it means to say that "... both countries must produce *different* goods ..." or that "... both countries must produce the *same* good, ..." (p. 244). Why *must* they do either of these? He further says that "... if both countries must produce the *same* good, there is a value of $|k^a - k^b|$ above which that good must be exported by the country abundant in its intensively used factor." But what happens, as seems likely, if the "same good" turns out to be the intermediate one?

do not seem to have advanced much beyond my statement of the theorem.

In conclusion, it should be pointed out that the discussion surrounding Stewart's Figure 3 is incorrect, for since the origins for the two

countries are not the same, the quantities produced for the country producing at R cannot be shown on this diagram, and the quantity comparisons he draws are therefore inappropriate.

IN MEMORIAM

JACOB VINER

1892–1970

Jacob Viner, one of the great teachers and scholars of economics of our time and all times and former president (1939) of the American Economic Association, died in Princeton, New Jersey on September 12, 1970 at the age of 78.

Born in Montreal, Canada on May 3, 1892, Professor Viner attended McGill University (B. A., 1914) and then did his graduate work at Harvard University (M.A., 1915; Ph.D., 1922). His doctoral dissertation, written under Professor Frank Taussig, received the David Ames Wells Prize, and was published in 1924 as *Canada's Balance of International Indebtedness*. This book, which quickly established his reputation as an economic theorist of the first rank, not only extended the pure theory of international adjustment but was also a pioneer venture in combining rigorous theory with the thorough empirical testing of theoretical propositions.

In 1916, Viner joined the faculty of economics of the University of Chicago, where he remained until 1945, when he moved to Princeton University. He retired from teaching in 1960, but continued active intellectual work as an Emeritus Professor at Princeton and as a permanent member of the Institute for Advanced Study.

The American Economic Association awarded him its highest honor, the Francis A. Walker Medal, in 1962. Thirteen universities in the United States and abroad recognized his professional achievements with honorary degrees.

Viner was active in public affairs, particularly from 1934 to 1939, when he was special assistant to the Secretary of the Treasury, and when he played an important role in developing a research staff at the Treasury and in shaping national economic policy. In addition, he was special adviser to the U.S. Tariff Commission and U.S. Shipping Board during World War I; consultant to the De-

partment of State 1943 to 1952; consultant to the Board of Governors of the Federal Reserve System; member of the Board of Directors of the National Bureau of Economic Research.

Yet these public activities were very much a side line. Viner's primary interests and activities were academic. Like his own mentor, F. W. Taussig, Viner was a great teacher. His basic course in Price Theory at the University of Chicago, which many students took in their first quarter of graduate study, was a deep intellectual experience. Some found it forbidding and developed a fear of the man and the subject. But to the abler students, it opened a new world. It gave them a feeling for the subtlety, power, and appeal of pure economic theory. In Viner's hands, economic theory was not a set of formal abstract propositions; it was a set of tools, to be constructed with care and the utmost attention to logical rigor, but to be judged primarily by its usefulness in understanding and interpreting important economic phenomena. He presented economic theory as, in Marshall's words, "an engine of analysis." And he presented it with verve and color, making it an exciting and controversial subject. He had few peers for quickness of mind and tongue or ability to grasp new ideas or to spot and expose fallacies.

Like Taussig also, he was a great editor. He edited or coedited the *Journal of Political Economy* for 18 years, from 1928 to 1946, keeping it in the forefront of the professional economic journals of the world. He set and maintained the very highest standards of integrity in editing, and displayed a catholicity and tolerance that made the *JPE* a widely ranging journal containing contributions embodying different approaches and reflecting different schools of thought.

Viner's scholarly influence was exerted in many areas: In price theory, where his article

on cost curves has been justly famous as a major improvement in Marshall's analysis of the relation between costs and supply; in the history of economic thought, where his incisive review of the enormous mercantilist pamphlet literature led to the rehabilitation of Adam Smith's judgment on mercantilism as a theory; in studies of the balance of power and the economic aspect of imperialism, which stimulated extensive further work by his students as well as other scholars.

His major scholarly impact, however, was unquestionably in international trade. In addition to his doctoral thesis, he published a number of seminal works, including *Dumping* (1922), *Studies in the Theory of Inter-*

national Trade (1937), and *Customs Unions* (1950). His *Studies* combined his special interest in trade with his deep and abiding interest in the history of economic thought. It remains the definitive work on many aspects of the development of both monetary theory and international trade theory. His small book on *Customs Unions* introduced a fresh analysis of an ancient subject that has dominated further discussion of the problem.

Viner was a man of wide culture and interests, a fascinating conversationalist, warm friend, devoted father and husband, an ever-ready source of advice, and help to a legion of present and former students, colleagues and friends.

ERRATA

The Optimum Lifetime Distribution of Consumption Expenditures: A Correction

By LESTER C. THURLOW

In an iteration of comments and responses, the September 1970 *Review* inadvertently published my original response to the comments of Brian Motley and Samuel A. Morley on "The Optimum Lifetime Distribution of Consumption Expenditures" rather than my final response (pp. 744-45). The last three paragraphs of the published reply are incorrect. They should be replaced by the following three paragraphs.

If the quantitative magnitudes of the

estimates are to be explained by this factor, however, there must be a very peculiar distribution of lifetime incomes over different zero-saver age groups. The lifetime income of a zero-saver man 60 years old would have to be just half of the lifetime income of a zero-saver man 40 years old if this factor were to explain savings behavior by itself. (Assuming that economic growth leads to a real rate of growth of 2 percent per year in family income, the average lifetime income of a 40 year old is 44 percent higher than that of a 60 year old.) At the other end of the scale, the lifetime income of a zero-saver man 20 years old would have to be less than the lifetime income of a zero-saver man 40 years old to explain the estimated results. I can swallow the former but not the latter.

Given identical and homothetic preferences, I agree with Motley and Morley that equation (1) is the correct estimating equation. *It is exactly the equation which I used.* The problem is to measure the general average lifetime income (\bar{W}) and the average lifetime income of the zero saver (W^*). In

my paper the definitions were as follows:

$$(1) \quad \bar{W} = \sum_{t=1}^T \bar{Y}_t$$

$$(2) \quad W^* = \sum_{t=1}^T Y_t^* = \sum_{t=1}^T C_t^*$$

where:

$Y_t^* = C_t^*$ for the zero-saver

\bar{Y}_t = income of average t -year old household.

I also agree that both are at best only approximations to actual lifetime incomes. Fortunately some of the errors in such an approximation may disappear since the same types of biases are apt to appear in both \bar{W} and W^* , and only the ratio of \bar{W} to W^* is used in the analysis.

I find the Motley-Morley method of calculating optimum life-time incomes interesting, but I am puzzled as to why their model becomes the Thurow technique when it yields incorrect results. They show that equation (7) yields a convex path when the normal income path is concave and the true optimal path is postulated as constant. I agree with the conclusions, but equation (7) is not the estimating equation which I used. It is even labelled an alternative model. Thus, I can hardly consider the incorrect results a critique of my estimating equation.

A Growth Model of International Direct Investment

By HANS BREMS

In the June, 1970 issue of this *Review* on p. 328, I assumed entrepreneurs to ignore their own influence upon national money income Y_i and hence upon the volume of savings $(1-c_i)Z_i$; they were allocating between parent firm and foreign subsidiary. This assumption was unnecessary. Use (92) through (95) and (100) through (103) to see that $g_{P11} + g_{X11} = G + g_{P21} + g_{X21}$, hence

$$Z_1(\tau) = [r_1 - (g_{P11} + g_{X11})]\xi_1(\tau)$$

Insert this into (60) written for time τ ; then take the partial derivative of (60) with respect to $S_{11}(\tau)$

$$\frac{\partial Y_1(\tau)}{\partial S_{11}(\tau)} = [r_1 - (g_{P11} + g_{X11})] \frac{\partial \xi_1(\tau)}{\partial S_{11}(\tau)}$$

Consequently, maximizing $\xi_1(\tau)$ by taking the partial derivative $\partial \xi_1(\tau)/\partial S_{11}(\tau)$ and setting it equal to zero also makes the partial derivative $\partial Y_1(\tau)/\partial S_{11}(\tau)$ zero.

Thus all solutions of my model hold whether or not entrepreneurs ignore their own influence upon national money income and savings.

NOTES

An amendment to the bylaws adopted by mail ballot last year provides for nomination by petition to the offices of President-elect, Vice President, and member of the Executive Committee. To be valid, petitions must reach the Secretary by August 1. Petitions for President-elect must be signed by 6 percent of the members; for other offices by 4 percent of the members. The exact text of the new bylaw will be published in the May issue of this *Review*.

The names of nominees selected by the Nominating Committee will be published in the June issue of this *Review*. Members wishing to have this information earlier may contact the Secretary by mail or telephone at any time after April 1. The Secretary will make every effort to facilitate efforts to obtain the requisite number of signatures on petitions. In particular, address labels for all members will be furnished at cost.

The Economics Institute is sponsored by the American Economic Association for foreign students of economics and agricultural economics beginning graduate work in the United States.

The 14th session is to be held from June 16 to August 18, 1971 at the University of Colorado. Information and application forms may be obtained from the Director of the Economics Institute, University of Colorado, Boulder, Colorado 80302.

The American Economic Association Nominating Committee for 1971 was listed on page 1021 of the December 1970 *Review*. Karl Gregory, Oakland University should be added to this list. The omission of his name was an error.

The American-Scandinavian Foundation, the Council for European Studies, and the Society for the Advancement of Scandinavian Study will cosponsor a one-day symposium/convention devoted to "Social Science Research in Scandinavia." The symposium, to be held in conjunction with the annual meeting of S.A.S.S. in Lexington, Kentucky next May, will focus on the problems and methodology of scholarly research in the Nordic countries in the fields of history, political science, economics, sociology, and other closely related fields.

The academic director of the symposium will be Stephen Bland, associate professor of political science, University of Pittsburgh and executive director, Council for European Studies. The administrative director is Gene G. Gage, executive secretary of The American-Scandinavian Foundation and secretary-treasurer of The Society for the Advancement of Scandinavian Study. For further information, contact: Mr. Gage,

ASF, 127 East 73rd Street, New York, N.Y. 10021, Telephone 212+TR9-9779.

The 1971 annual meeting of the New York State Economics Association (NYSEA) will be held April 16-17 at Rensselaer Polytechnic Institute, Troy, New York. Officers of NYSEA for 1970-71 are: President: Edwin J. Holstein, Rensselaer Polytechnic Institute; Vice President: David A. Martin, State University College at Geneseo; Secretary-Treasurer: Nancy R. Auster, State University of New York Agricultural and Technical College at Canton; Editor: Daniel Feinberg, New York City Community College.

New data for researchers on labor force behavior and work attitudes: First results of the early surveys made for a five-year study of the labor market experience, characteristics, and work attitudes of four groups of 5,000 people each are being made available on data tapes as well as in print. Data for the four groups: men 45-59 years old; women 30-44 years old; and young men and women 14-24 years old; come from annual surveys funded by the Labor Department's Manpower Administration. The final surveys will be conducted in 1973. The surveys, beginning in 1966 for the two groups of men, 1967 for the older women, and 1968 for the young women, cover: Current employment, work history, education and training, migration, earnings, income, family status, health, work attitudes; plus information of special significance for each group. For example, retirement plans of older men, and the costs of child care and transportation for employed older women.

Data tapes may be obtained at cost through the Chief of the Demographic Surveys Division, Bureau of the Census, U.S. Department of Commerce, Washington, D.C. 20233. Tape specifications were published in late 1970. The tapes will not identify individuals, in line with the Census Bureau's policy of confidentiality.

Printed reports may be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Reports now available cover the first two studies of older men, the first study of young men, and the first study of older women. Order as Manpower Research Monographs Nos. 15-I, 15-II, 16-I, and 22-I, respectively.

A visiting lecturer program in statistics has been organized for the eighth successive year. The program is sponsored jointly by the principal statistical organizations in the United States, the American Statistical Association, the Biometric Society, and the Institute of Mathematical Statistics. The National Science Foundation provides partial financial support. Leading teachers and research workers in statistics from universities, industry, and government, have agreed to participate as lecturers. Lecture topics include subjects in experimen-

tal and theoretical statistics, as well as in such related areas as probability theory, information theory, and stochastic models in the physical, biological, and social sciences.

The purpose of the program is to provide information to students and college faculty about the nature and scope of modern statistics, and to provide advice about careers, graduate study, and college curricula in statistics. Inquiries should be addressed to: Visiting Lecturer Program in Statistics, Department of Statistics, Southern Methodist University, Dallas, Texas 75222.

The National Science Foundation will sponsor the following College Teacher Summer Programs for 1971:

For teachers of economics: June 7–July 30. Mathematics for economics, economic theory (price and national income), managerial economics, and stabilization and growth policies. Dr. William H. Wesson, Jr., department of economics, University of South Carolina, Columbia, South Carolina 29208.

June 27–July 24: Urban economic growth and structure, human resources in urban economies, and the public sector of urban economies. Dr. Henry Levin, department of economics, Stanford University, Stanford, California 94305.

June 21–August 13: Micro- and macroeconomic analysis and its relationship to the principles courses; workshop on problems of junior college teachers of introductory economics. Dr. Jay W. Wiley, department of economics, Purdue University, Lafayette, Indiana 47907.

For teachers of introductory economics: June 7–July 16. Current issues of economic policy; and the development and teaching of a policy oriented, basic course in economics. Dr. Ewing P. Shahan, department of economics and business administration, Vanderbilt University, Nashville, Tennessee 37203.

The 1971 annual meeting of the Association of Social and Behavioral Scientists (formerly the Association of Social Science Teachers) will be held April 22 and 23 at the Holiday Inn Downtown, Montgomery, Alabama. The theme is "Environment: The Ghetto and Beyond." Program chairman is Russell Stockard, and inquiries may be addressed directly to him at the department of geography, Southern University at New Orleans, New Orleans, Louisiana.

The Public Policy Program at Harvard University is conducting an open competition to secure case studies of the application of economic and other analytic methods to real public policy problems. All completed works not previously published are eligible. Studies that have been employed in a policy-making context are particularly sought. There is no limit on length; extensive original data may be included.

Submissions will be reviewed promptly by a committee under the direction of Professors Raiffa, Schelling, and Zeckhauser. Selected entries will receive informal publication by the Public Policy Program. They

will be used in our teaching program and will be distributed to other interested readers. The authors of selected entries will maintain all rights and copyright privileges, and will be awarded a \$100 honorarium.

Manuscripts will be returned to authors on request. For curriculum planning purposes, it is important that entries be received not later than June 30, 1971. They should be sent to R. Zeckhauser, Littauer Center M-41, Harvard University, Cambridge, Massachusetts 02138.

Summer NSF Institute in management science and operations research for college professors in management science, operations research, applied mathematics, economics, health and hospital administration to be held June 14 through July 16, 1971 at the University of Colorado, Boulder. Participants receive stipend and travel expenses.

For further information, correspond with: Dr. Donald R. Plane, Co-Director, Summer Institute, Division of Management Science, Business Building, University of Colorado, Boulder, Colorado 80302.

Deaths

Elmer C. Bratt, professor of economics emeritus, Lehigh University, Nov. 9, 1970.

Alfred H. Conrad, professor of economics, City College of New York, Oct. 18, 1970.

Howard E. Dubner, assistant professor of economics, University of Miami, Oct. 9, 1970.

Everett D. Hawkins, professor of economics, University of Wisconsin, Aug. 31, 1970.

Milton S. Heath, professor of economics emeritus, University of North Carolina, Chapel Hill, Sept. 8, 1970.

John B. Lansing, chairman, department of economics, University of Michigan, Sept. 8, 1970.

Ben S. Seligman, professor of economics and director, Labor Relations and Research Center, University of Massachusetts, Oct. 23, 1970.

Retirements

C. Richard Creek, department of economics, Colorado State University, Sept. 1970.

John M. Frikart, professor of economics, University of Arizona, June 1971.

Kent T. Healy, department of economics, Yale University, July 1, 1970.

Ross Milner, professor of agricultural economics, Ohio State University, July 31, 1970.

Arthur H. Reede, professor of economics, Pennsylvania State University, Sept. 1970.

Frederick G. Reuss, department of economics, Goucher College.

Josef Soudek, professor of economics, City University of New York, Queens College, Sept. 1970.

Visiting Foreign Scholars

Knolly Barnes, University of West Indies: visiting scholar, department of economics, University of Nebraska.

Michael Bruno, Hebrew University: visiting professor, department of economics, Massachusetts Institute of Technology.

Frits J. de Jong, University of Groningen: National Science Foundation Senior Foreign Scientist Fellowship, Florida State University, 1970-71.

Gunnar Floystad, Norwegian School of Economics and Business Administration: visiting scholar, department of economics, University of Michigan, 1970-71.

Frank H. Hahn, London School of Economics: visiting professor, department of economics, Massachusetts Institute of Technology.

Roy Harrod, Oxford University: visiting professor of economics, University of Maryland, spring 1971.

Terence W. Hutchison, University of Birmingham: visiting professor of economics, Dalhousie University, fall 1970.

Sven-Ake Johansson, Royal Institute of Technology, Stockholm, Sweden: visiting scholar, Graduate School of Business Administration, University of California, Los Angeles, Oct. 1970.

S. S. Johl, Punjab Agricultural University, India: visiting professor of economics, Ohio State University, Aug. 1970-Aug. 71.

Yoav Kislev, Hebrew University of Jerusalem: visiting research associate, department of economics, Yale University, 1970-71.

James Mirrlees, Nuffield College, England: visiting professor, department of economics, Massachusetts Institute of Technology.

Carl J. Norstrom, Norwegian School of Economics and Business Administration: visiting scholar, department of economics, University of Michigan, 1970-71.

Seiji Sakurai, Hokkaido University: visiting scholar, department of agricultural economics, Cornell University.

Hirofuma Shibata, York University, England: visiting associate professor of economics, University of Maryland, 1970-71.

Yoshio Shimizu, Konan University, Kobe, Japan: visiting scholar, department of economics, University of Colorado, 1970-71.

Baran M. Tuncer, University of Ankara: visiting research associate, department of economics, Yale University, 1970-71.

A. J. W. van de Gevel, Tilburg University: visiting assistant professor of economics, Northern Illinois University, 1970-71.

Henricus P. A. van Roosmalen, Institute of Social Studies: visiting associate professor of business administration, University of New Hampshire, fall 1971.

Promotions

Richard J. Agnello: assistant professor of business and economics, University of Delaware, June 1970.

R. G. Akkihal: associate professor of economics, Marshall University.

Thaine H. Allison, Jr.: assistant professor of economics, Central Washington State College.

Nancy R. Auster: associate professor of economics, State University of New York, Canton.

Robin Barlow: professor of economics, University of Michigan.

Merrill J. Bateman: professor of economics, Brigham Young University.

Paul T. Bechtol: professor of economics, Colorado College.

Robert A. Behren: associate professor of economics, Brooklyn College.

William C. Bonifield: associate professor of economics, Wabash College.

Eleutherios N. Botsas: associate professor of economics and management, Oakland University, July 1970.

Patricia F. Bowers: associate professor of economics, Brooklyn College.

Charles W. Bullard: professor of economics, University of North Dakota.

Winston Chang: associate professor of economics, State University of New York, Buffalo, Sept. 1970.

Allan R. Cchen: associate professor of business administration, Whittemore School of Business and Economics, University of New Hampshire.

Jerry L. Dake: associate professor of corporate finance, College of Industrial Management, Georgia Institute of Technology.

Robert H. Deans: associate professor of economics, Temple University.

Lawrence P. Donnelley: assistant professor of business and economics, University of Delaware, June 1970.

Robert B. Ekelund: associate professor of economics, Texas A&M University.

Edward G. Emerling: professor of economics, St. Bonaventure University.

John T. Etheridge: associate professor of industrial management, College of Industrial Management, Georgia Institute of Technology.

Irwin Feller: associate professor of economics, Pennsylvania State University.

Jay D. Forsyth: assistant professor of business administration, Central Washington State College.

Wolfgang W. Franz: assistant professor of economics, Central Washington State College.

A. Myrick Freeman III: associate professor of economics, Bowdoin College, fall 1970.

Bruce L. Gensemer: associate professor of economics, Kenyon College.

Richard A. Goodman: associate professor, department of business administration, Graduate School of Business Administration, University of California, Los Angeles.

Richard L. Gordon: professor of mineral economics, Pennsylvania State University, July 1, 1970.

William P. Gramm: associate professor of economics, Texas A&M University.

Bennett Harrison: assistant professor of economics, University of Maryland, Jan. 1971.

Robert E. Hicks: associate professor, department of economics and finance, Florida Technological University.

C. Russell Hill: assistant professor of economics, University of Michigan.

James M. Holmes: associate professor of economics, State University of New York, Buffalo, Sept. 1970.

Teh-wei Hu: associate professor of economics, Pennsylvania State University, July 1970.

John R. Kaatz: associate professor of economics, College of Industrial Management, Georgia Institute of Technology.

Alvin K. Klevorick: associate professor of economics, Yale University, July 1970.

A. J. Kondonassis: professor of economics, University of Oklahoma, Sept. 1970.

Ronald S. Koot: associate professor of quantitative business analysis, Pennsylvania State University, July 1970.

Iwan S. Koropecykj: professor of economics, Temple University.

Chung H. Lee: associate professor of economics, Miami University.

Albert M. Levenson: professor of economics, Queens College, City University of New York.

Daniel Lipsky: professor of economics, Brooklyn College.

Wesley Long: associate professor of economics, Pennsylvania State University.

Robert M. Lovejoy: associate professor of economics, State University of New York, Binghamton.

John J. McGowan: associate professor of economics, Yale University, July 1, 1970.

Graeme H. McKechnie: associate professor of economics, York University.

William J. McKenna: professor of economics, Temple University.

Donald McLeod: associate professor of economics, Temple University.

James B. MacQueen: professor, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Fredric C. Menz: assistant professor of economics, Temple University.

Roger N. Millen: assistant professor, Graduate School of Business, Columbia University, July 1, 1970.

Charles G. Moore: assistant professor of economics, Brooklyn College.

Dennis C. Mueller: associate professor, department of economics, Cornell University, July 1, 1970.

Eugene A. Myers: professor of economics, Pennsylvania State University, July 1970.

William D. Nordhaus: associate professor of economics, Yale University, July 1, 1970.

Carl Nordstrom: professor of economics, Brooklyn College.

J. Randolph Norsworthy: associate professor of economics, Temple University.

Patrick R. O'Shaughnessy: associate professor of business administration, Central Washington State College.

Lewis J. Perl: assistant professor, department of labor economics and income security, New York State School of Industrial and Labor Relations, Cornell University.

Barry W. Poulson: associate professor of economics, University of Colorado.

Arnold H. Raphaelson: professor of economics, Temple University.

T. Ross Reeve: assistant professor of economics, Central Washington State College.

Richard D. Reimer: professor of economics, College of Wooster, Sept. 1970.

Stephen A. Resnick: associate professor of economics, Yale University, July 1, 1970.

Richard Rosenberg: assistant professor of economics, Pennsylvania State University, Jan. 1970.

George H. K. Schenck: associate professor of mineral economics, Pennsylvania State University, July 1, 1970.

Harvey Schwartz: associate professor of economics, York University.

Richard A. Seese: assistant professor of economics, John Carroll University.

Paul Seidenstat: associate professor of economics, Temple University.

Merrill K. Sharp: assistant professor, department of economics, Iowa State University.

Philip Sheinwold: professor of economics, Brooklyn College.

Mitchell Stengel: assistant professor, department of economics and Center for Urban Affairs, Michigan State University, Sept. 1, 1970.

Samuel S. Stewart: assistant professor, Graduate School of Business, Columbia University, July 1, 1970.

Joseph E. Stiglitz: professor of economics, Yale University, July 1, 1970.

Venkataraman Sundararajan: assistant professor of economics, University College, New York University.

Richard J. Trethewey: assistant professor of economics, Kenyon College.

Joan G. Walters: associate professor of economics, Fairfield University.

Martin L. Weitzman: associate professor of economics, Yale University, July 1, 1970.

Katherine M. West: associate professor of economics, Brooklyn College.

Administrative Appointments

John W. Allen: chairman, department of economics, Texas A&M University.

Robert Andrews: assistant dean, Graduate School of Business Administration, University of California, Los Angeles, Oct. 1970.

Robert L. Aronson: chairman, department of labor economics and income security, New York State School of Industrial and Labor Relations, Cornell University, July 1970.

D. A. L. Auld: acting chairman, department of economics, College of Social Science, University of Guelph, July 1, 1970.

Morton S. Baratz: director, Boston University Urban Institute.

Thomas A. Bausch: assistant dean, School of Business, John Carroll University.

Charles E. Bishop, North Carolina University: chancellor, University of Maryland, fall 1970.

David B. Brooks, U.S. Bureau of Mines: head, economic research section, Mineral Resource Branch, Department of Energy, Mines and Resources, Ottawa.

Elwood S. Buffa: associate dean, Graduate School of Business Administration, University of California, Los Angeles, July 1970.

Herbert A. Chesler: administrative officer, department of economics, University of Pittsburgh.

Young-lob Chung: chairman, department of economics, Eastern Michigan University.

Lawrence P. Cole: assistant dean, Whittemore School of Business and Economics, University of New Hampshire.

George G. Dawson: director of research and publications, Joint Council on Economic Education; managing editor, *Journal of Economic Education*, Sept. 1970.

Lynn E. Dellenbarger, Jr.: director of graduate programs in business administration and industrial relations, College of Business and Economics, West Virginia University.

John Dorsey: vice-chancellor for Business Affairs, University of Maryland, fall 1970.

James S. Earley: dean, College of Social and Behavioral Sciences, University of California, Riverside, July 1970.

Leo Fishman: chairman, department of economics, College of Business and Economics, West Virginia University.

Michael Gort: acting chairman, department of economics, State University of New York, Buffalo, 1970-71.

R. Earl Green: associate dean, College of Industrial Management, Georgia Institute of Technology.

John D. Guilfoil: director of undergraduate studies, department of economics, New York University, Washington Square.

Bernard Hall: vice-president and provost, Kent State University.

John R. Haskell: assistant dean, Whittemore School of Business and Economics, University of New Hampshire.

Robert E. Hicks: acting chairman, department of economics and finance, Florida Technological University.

Alfred E. Hofflander: vice-chairman, department of business administration, Graduate School of Business Administration, University of California, Los Angeles, Oct. 1, 1970.

O. Henry Hoversten: head, department of business administration and economics, Capital University.

Thomas S. Isaack: chairman, department of business administration, College of Business and Economics, West Virginia University.

Harald R. Jensen: acting head, department of agricultural and applied economics, University of Minnesota.

George D. Johnson: chairman, accounting and quantitative methods department, Chico State College.

Mark J. Kasoff: chairman, department of economics, Antioch College.

J. David Lages: chairman, department of economics, Southwest Missouri State College.

Charles N. Lanier: acting chairman, department of economics, University of Delaware, 1970-71.

Albert M. Levenson: associate dean, Faculty for the

Social Sciences, Queens College, City University of New York.

W. Clair Lillard: director, International Programs, Central Washington State College.

Kenneth McLennan: chairman, department of economics, Temple University.

Patrick Mann: director, graduate programs in economics, College of Business and Economics, West Virginia University.

Laurence J. Mauer: acting chairman, department of economics, Northern Illinois University.

Frederic Meyers: vice-chairman, department of business administration, Graduate School of Business Administration, University of California, Los Angeles, Oct. 1, 1970.

John R. Morris: associate chairman, department of economics, University of Colorado.

John H. Niedercorn: chairman, department of economics, University of Southern California.

Martin Plotnik: acting chairman, department of economics, Slippery Rock State College, Jan. 1971.

Roger L. Ranson: chairman, department of economics, University of California, Riverside, July 1970.

William R. Reilley: chairman, department of economics and business administration, Norwich University, July 1970.

Warren L. Smith: acting chairman, department of economics, University of Michigan.

Joseph M. Thorson: associate dean, School of Social and Behavioral Sciences, West Chester State College.

Joseph L. Tryon: chairman, department of economics, Georgetown University, summer 1971.

Hugh G. Wales, University of Illinois: acting head, department of management, College of Business Administration, Roosevelt University.

Laszlo Zsoldos: acting dean, College of Business and Economics, University of Delaware, 1970-71.

Appointments

Jan P. Acton, Harvard University: staff member, economics department, The RAND Corporation, Nov. 1970.

Arjun Adlakha: research demographer, Food Research Institute, Stanford University.

Hamilton Alexander, Old Dominion University: assistant professor of economics, Virginia Commonwealth University.

Robert Ante: assistant professor of geography, department of economics, Queens College, City University of New York, 1970-71.

Sven W. Arndt, University of California, Santa Cruz: visiting assistant professor, Food Research Institute, Stanford University, spring 1971.

Enrique Arzac: assistant professor, Graduate School of Business, Columbia University, Jan. 1, 1971.

E. Dean Baldwin, University of Illinois: assistant professor of economics, Miami University.

Joseph L. Balintfy: professor of general business and finance; department of industrial engineering, University of Massachusetts, Sept. 1970.

Milton J. Bass, University of Michigan: lecturer, de-

partment of economics, Queens College, City University of New York, 1970-71.

Henry B. R. Beale: instructor, department of economics, Georgetown University, 1970-71.

Larry G. Beall, Erskine College: assistant professor, department of economics, Virginia Commonwealth University.

Burley V. Becholdt: assistant professor of economics, Northern Illinois University.

William S. Becker, Louisiana State University: assistant professor, department of economics, Colorado College.

Larry D. Bedford: instructor, department of economics, Iowa State University.

Sanford A. Belden: assistant professor, department of agricultural economics, Cornell University.

Herman A. Berliner, City University of New York: assistant professor of economics, Hofstra University, fall 1970.

George W. Betz, University of Singapore, Malaya: associate professor of economic development, Whittemore School of Business and Economics, University of New Hampshire.

Edward R. Bleau: instructor, department of economics, Marshall University, 1970-71.

Sam H. Book: assistant professor, faculty of administrative studies, York University, Sept. 1970.

J. Patrick Bovino, Pennsylvania State University: instructor of business administration, Whittemore School of Business and Economics, University of New Hampshire, fall 1971.

Leonard R. Boyer, Kent State University: instructor of economics, Old Dominion University.

Colin I. Bradford, Jr.: senior economist, CIAP, Department of Economic Affairs, Organization of American States, Washington.

Eric Brucker, Southern Illinois University: assistant professor of business and economics, University of Delaware, Sept. 1970.

John Buttrick: professor, department of economics, York University, June 1970.

William P. Butz, University of Chicago: staff member, economics department, The RAND Corporation, Sept. 1970.

Jann W. Carpenter: associate professor of business administration, Central Washington State College.

Johanna R. Cavanaugh: lecturer, accounting department, University of Nevada, 1970-71.

Fikret Ceyhan, Wayne State University: assistant professor of economics, University of North Dakota, 1970-71.

Chau-nan Chen: assistant professor of economics, Northern Illinois University.

C. Mark Choate: assistant professor of general business and finance, University of Massachusetts, Sept. 1, 1970.

Ronald Choy, University of California, Berkeley: staff member, economics department, The New York City RAND Institute, July 1970.

David S. C. Chu, Yale University: staff member, economics department, The RAND Corporation, Oct. 1970.

Donald J. Cocheba: assistant professor of business administration, Central Washington State College.

Patricia A. Coffey: instructor, department of economics, Iowa State University.

Neal P. Cohen, University of Wisconsin: assistant professor of economics, Eastern Michigan University.

Richard V. L. Cooper, University of Chicago: staff member, economics department, The RAND Corporation, Jan. 1971.

Philip G. Cotterill, University of Illinois, Chicago Circle: assistant professor of economics, Miami University.

Larry Cox, Kansas State University: instructor of economics, Southwest Missouri State College.

James F. Crook: assistant professor, School of Business Administration, Winthrop College, fall 1970.

Alvin M. Cruze: visiting lecturer in economics, University of North Carolina, Chapel Hill.

Robert Daniels, University of Lancaster: assistant professor, department of economics, Case Western Reserve University.

Rachel Dardis, Cornell University: lecturer in economics, associate professor of home economics, University of Maryland, fall 1970.

Paul G. Darling, Bowdoin College: visiting professor of economics, McGill University, spring 1971.

William W. Davis, University of Kentucky: instructor, department of economics, Western Kentucky University, 1970-71.

Michael E. dePrano, University of Southern California: visiting professor of economics, Texas A&M University.

John Despres, University of California, Berkeley: staff member, economics department, The RAND Corporation, Sept. 1970.

James Doane, Northeastern University: assistant professor, department of economics, University of Maine.

Michael R. Dohan, California Institute of Technology: assistant professor of economics, Queens College, City University of New York, Feb. 1971.

Peter F. Drucker, New York University: professor of business economics, Claremont Graduate School, spring 1971.

Donald H. Ebbeler: assistant professor of economics, College of Industrial Management, Georgia Institute of Technology.

Linda N. Edwards, Columbia University: lecturer, department of economics, Queens College, City University of New York, 1970-71.

Donald Ellenstine, Baker University: associate professor of economics, Eisenhower College.

Doyle A. Eiler: assistant professor, department of agricultural economics, Cornell University.

Ghazi T. Farah, Wisconsin State University: assistant professor of economics, Florida Technological University.

John D. Ferguson, Brown University: assistant professor of economics, Miami University.

Sherwood M. Fine: visiting professor of economics, Washington and Lee University.

James J. Fralick: assistant professor, department of economics, Fordham University, 1970-71.

James H. Gapinski, University of New York, Buffalo: assistant professor, department of economics, Florida State University.

Michael H. Gieske: research associate, department of economics, Iowa State University.

Roy F. Gilbert, Michigan State University: visiting assistant professor of economics, Texas A&M University.

Frederic Glantz, Syracuse University: instructor of economics, Temple University.

Martin L. Gosman: assistant professor of accounting, University of Massachusetts, Sept. 1, 1970.

Harold G. Halcrow, University of Illinois: visiting professor, Food Research Institute, Stanford University, winter quarter 1970-71.

Robert E. Hall, University of California, Berkeley: associate professor, department of economics, Massachusetts Institute of Technology.

Stanley H. Hargrove: research associate, department of economics, Iowa State University.

Michael J. Hartley, Duke University: assistant professor of economics, State University of New York, Buffalo, Sept. 1970.

Joseph A. Hasson: visiting professor of economics, Acadia University, 1970-71.

Walter E. Hecox, U.S. Military Academy: assistant professor, department of economics, Colorado College.

James J. Heilbrun: associate professor, department of economics, Fordham University, 1970-71.

Julius Held: assistant professor of economics, Brooklyn College.

Richard Hellman, Small Business Administration: visiting professor of economics, University of Rhode Island, fall 1970.

George R. Henderson: instructor of economics, Miami University.

Louis Henry, Notre Dame University: assistant professor of economics, Old Dominion University.

Larry M. Hersh, Harvard University: lecturer, department of economics, Queens College, City University of New York, 1970-71.

S. Hugh High, Mississippi State College: instructor, department of economics, North Texas State University.

Roger B. Hill, University of Georgia: professor of economics, University of North Carolina, Wilmington.

George E. Hoffer, University of Virginia: assistant professor, department of economics, Virginia Commonwealth University.

Joe R. Hulett, Iowa State University: assistant professor of economics, Texas A&M University.

John A. Hutcheson: visiting assistant professor of economics, Queen's University.

Milton A. Iyoha, Yale University: assistant professor of economics, State University of New York, Buffalo, Sept. 1970.

William Jaffe: professor, department of economics, York University, June 1970.

John B. Kaye, George Washington University: lecturer, management department, University of Nevada, 1970-71.

William M. Kempey, Hofstra University: teaching associate, department of economics, Queens College, City University of New York, 1970-71.

A. Thomas King, Yale University: research assistant, Bureau of Business and Economic Research; lecturer in economics, University of Maryland, fall 1970.

Walter J. Klages, York University: associate professor of economics, Florida Technological University.

Tetteh A. (Ben) Kofi: acting assistant professor, Food Research Institute, Stanford University.

Joseph Kowalski, Wayne State University: instructor of economics, Temple University.

Theodore J. Kreps: visiting professor of economics, Central Washington State College.

Prem Kumar: assistant professor of general business and finance, University of Massachusetts, Sept. 1, 1970.

Eddy L. LaDue: assistant professor, department of agricultural economics, Cornell University.

Donald Larson, Michigan State University: assistant professor of economics, Ohio State University, Oct. 1, 1970.

Robert J. Latham: assistant professor of economics, Pennsylvania State University, Sept. 1970.

Wendy Lee, Northwestern University: assistant professor of economics, Texas A&M University.

Wayne E. Leininger: instructor in accounting, University of Massachusetts, Sept. 1, 1970.

Charles R. Link: instructor of business and economics, University of Delaware, Sept. 1970.

Jay Lowden, Jr.: instructor, division of business and economics, Nasson College, Sept. 1970.

Mark A. Lutz: assistant professor, department of economics, University of Maine.

Charles E. McConnel, Occidental College: assistant professor of economics, University of Texas, El Paso, 1970-71.

Majorie McElroy: assistant professor, department of economics, Duke University, 1970-71.

W. Terrence McGrath, University of Southern California: visiting assistant professor of economics, University of Maryland, 1970-71.

Jacob Merrieweather: associate in business, Graduate School of Business, Columbia University, Oct. 1, 1970.

Richard G. Milk: assistant professor of economics, College of Business Administration, Northeast Louisiana University.

Ira J. Miller: assistant professor, department of economics, Southern Methodist University, fall 1970.

L. Charles Miller, Jr., Haverford College: assistant professor of economics and director of internship program in urban studies, Yale University, July 1, 1970.

Stephen M. Miller: instructor, department of economics, University of Connecticut, Sept. 16, 1970.

William J. Moore, University of Oklahoma: assistant professor, department of economics, University of Houston.

Max Moszer, Bureau of Labor Statistics: professor of economics, Virginia Commonwealth University.

Roger Murray: professor of banking and finance, Graduate School of Business, Columbia University, Jan. 1, 1971.

Selma J. Mushkin: professor of economics, Georgetown University, 1970-71.

M. Ishaq Nadiri, Columbia University: professor of economics, New York University.

Lester A. Neidell: associate professor of marketing, College of Industrial Management, Georgia Institute of Technology.

Roger A. Norem: instructor, department of economics, Iowa State University.

Guy H. Orcutt, Urban Institute: professor of economics and urban studies, Yale University, July 1, 1970.

Alex Orden: professor of industrial management, College of Industrial Management, Georgia Institute of Technology.

Harry T. Oshima, University of Hawaii: visiting professor of economics, University of Pittsburgh, summer 1971.

David B. Pariser, Southern Illinois University: assistant professor of economics, University of North Dakota, 1970-71.

Robert J. Parsons: instructor, department of economics, Brigham Young University, 1970-71.

John Pisarkiewicz, Vanderbilt University: instructor, department of economics, Western Kentucky University, Aug. 1970.

Lon Polk: assistant professor of economics, Oakland University.

Daniel Primont, University of California, Santa Barbara: assistant professor of economics, Temple University.

Ronald E. Raikes: assistant professor, department of economics, Iowa State University.

Samuel D. Ramenofsky: instructor, department of economics, University of Missouri, fall 1970.

David E. Ramsett, University of Northern Iowa: associate professor of economics; director of North Dakota Center for Economic Education, University of North Dakota.

Suzanne E. Reid: assistant professor of economics, College of Business and Economics, West Virginia University.

Rudolf Rhomberg, International Monetary Fund: visiting professor of economics, University of Pittsburgh, fall/winter 1970-71.

Samuel B. Richmond: associate dean and professor, Graduate School of Business, Columbia University, Jan. 1, 1971.

M. Richard Roseman, University of Iowa: associate professor of economics, California State College, Los Angeles, Sept. 1970.

Bernard Rostker: economist, management sciences department, The RAND Corporation, Sept. 1970.

Timothy P. Roth, Texas A&M University: assistant professor, department of economics and finance, University of Texas, El Paso, 1970-71.

Dick L. Rottman, University of Missouri, Columbia: associate professor, University of Nevada, 1970-71.

Domenick Salvatore: instructor, department of economics, Forcham University, 1970-71.

Kehar S. Snagha, University of North Carolina, Charlotte: professor of economics, Old Dominion University.

Kazuo Sato, United Nations and Massachusetts Institute of Technology: professor of economics, State University of New York, Buffalo, Sept. 1970.

Linda Schumacher: instructor, department of economics, Hofstra University, fall 1970.

Abdelalem M. Sharshar, George Washington University: assistant professor, department of economics, Virginia Commonwealth University.

George W. Sheldon: instructor in general business and finance, University of Massachusetts, Sept. 1, 1970.

Robert Shishko, Yale University: staff member, economics department, The RAND Corporation, Sept. 1970.

Uma K. Sirivastava: research associate, department of economics, Iowa State University.

David J. Smyth, State University of New York, Buffalo: professor of economics, Claremont Graduate School, spring 1971.

Hugo H. Soll, Universidad Autonoma de Guadaluajara: assistant professor of economics, University of North Dakota.

John C. Soper: assistant professor, department of economics, Central Michigan University.

Philip E. Sorensen, University of California, Santa Barbara: associate professor, department of economics, Florida State University, Jan. 1971.

Brent W. Spaulding: instructor, department of economics, Iowa State University.

Samuel R. Stivers: assistant professor of industrial management, College of Industrial Management, Georgia Institute of Technology.

Donald E. Stone: associate professor of accounting, University of Massachusetts, Sept. 1, 1970.

Raymond S. Strangways, Georgia State University: professor of economics, Old Dominion University.

Martha Strayhorn, University of Wisconsin: lecturer in economics, University of Maryland, fall 1970.

Richard D. Teach: associate professor of marketing, College of Industrial Management, Georgia Institute of Technology.

William Thomas, U.S. Department of Commerce: assistant professor, department of economics, University of Houston.

G. Richard Thompson: assistant professor of economics, Florida Technological University.

Samuel L. Thorndike, Jr., University of Wisconsin, Milwaukee: assistant professor of economics, Bowdoin College, 1970-71.

Willard L. Thorp: professor of economics, University of Florida, Jan.-Mar. 1971.

John E. Tilton, Brookings Institution: assistant professor of economics, University of Maryland, fall 1970.

Subidey Togan, Johns Hopkins University: visiting assistant professor of economics, Texas A&M University.

Burke O. Trafton, American Metal Climax, Inc.: instructor of mineral economics, Pennsylvania State University.

Richard L. Tuley: instructor of economics, College of Industrial Management, Georgia Institute of Technology.

David J. Vail, Makerere University College, Uganda, Africa: assistant professor of economics, Bowdoin College, 1970-71.

Evangelina Vives: assistant professor of economics, Brooklyn College.

Robert P. Volyn, Upsala College: assistant professor of economics and business administration, Wagner College.

Gerald von Dohlen: assistant professor, Graduate School of Business, Columbia University, July 1, 1970.

Walter J. Wadycki, University of Illinois, Chicago Circle: assistant professor of economics and business analysis, Miami University.

Forrest E. Walters: associate professor, department of economics, Colorado State University.

Charles L. Weber: visiting assistant professor of economics, Valparaiso University, 1970-71.

James M. Weglin: lecturer in business administration, Central Washington State College.

David L. Weiss: lecturer in business administration, Central Washington State College.

Frederick J. Wells: visiting lecturer in economics, University of North Carolina, Chapel Hill, 1970-71.

Tom S. Witt: assistant professor of economics, College of Business and Economics, West Virginia University.

Dick R. Wittink: lecturer in business administration, Central Washington State College.

Edwin T. Wood: lecturer of business and economics, University of Delaware, Sept. 1970.

Robert A. Young: associate professor, department of economics, Colorado State University.

Jeffrey F. Zabler, University of Pennsylvania: assistant professor of economics, Wheaton College, Sept. 1970.

Leaves for Special Appointments

Dwight S. Brothers, Harvard University: program advisor in economic development and administration for East Africa, Ford Foundation, Sept. 1970-Aug. 1972.

Martin J. Davidson, North Texas State University: visiting professor, Bar-Ilan University, Israel, Jan. 1971.

William P. Dillingham, Florida State University: visiting professor, Fulbright Fellow, University of Rosario, Argentina.

Romesh K. Diwan, Rensselaer Polytechnic Institute: visiting professor, Graduate School of Economics, Washington University.

Thomas D. Duchesneau, University of Maine: staff economist, Bureau of Economics, Federal Trade Commission.

Isaiah Frank, Johns Hopkins University: executive director, President's Commission on International Trade and Investment Policy, 1970-71.

Walter Galenson, Cornell University: professor of American history and institutions, Cambridge University, England, 1970-71.

Jerome W. Hammond: University of Minnesota-Tunisia project under contract with the University of Minnesota and U.S. Agency for International Development.

Ernst W. Kuhn, University of Nebraska: Handelsschule, St. Gallen, Switzerland.

Donald Larson, Ohio State University: Institution Building Contract, Piracicaba, Brazil, Dec. 1970-72.

David F. Lean, Franklin and Marshall College: economist, Federal Trade Commission, 1970-71.

Scott R. Pearson, Food Research Institute, Stanford University: economist, President's Commission on International Trade and Investment Policy, 1970-71.

Malcolm J. Purvis, chief of University of Minnesota-Tunisia project under contract with the University of Minnesota and U.S. Agency for International Development.

Rex D. Rehnberg, Colorado State University: professor of agricultural economics, University of Nebraska-MASUA, program in Colombia, South America, 1970-72.

Lee P. Robbins, Franklin and Marshall College: assistant director, Urban Program of the Central Pennsylvania Consortium of Colleges, 1970-71.

Donald D. Rohdy, Colorado State University: consultant, dean of agriculture, Virginia Polytechnic Institute, 1970-71.

Larry E. Ruff, University of California, San Diego: U.S. Treasury Department, Office of Tax Analysis, Washington 1971-72.

Anthony E. Scaperlanda, Northern Illinois University: visiting professor of economics, Tilburg University, Tilburg, The Netherlands, 1970-71.

Wilson E. Schmidt, Virginia Polytechnic Institute and State University: deputy assistant secretary for research, Office of the Assistant Secretary for International Affairs, Department of Treasury, Washington.

Nathaniel E. Shechter, Old Dominion University: economic consultant to Bank of Israel, 1970-71.

Bertram Silverman, Hofstra University: research fellow, Yale University, summer and fall 1970.

Robert P. Strauss, University of North Carolina, Chapel Hill: Brookings Institution Economics Policy Fellow, 1970-71.

Samuel H. Talley, University of Maine: banking markets section, Division of Research and Statistics, Board of Governors of the Federal Reserve System.

David A. Walker, Pennsylvania State University: associate professor of quantitative business analysis, F.D.I.C., Washington, 1970-71.

Douglas W. Webbink, University of North Carolina, Chapel Hill: Brookings Institution Economics Policy Fellow, 1970-71.

J. Hugh Winn, Colorado State University: extension economist, Instituto Tecnológico Y De Estudios Superiores de Monterrey, Mexico, 1970-71.

Tae-Hee Yoon, Dominion Bureau of Statistics, Government of Canada: agricultural economist, International Bank for Reconstruction and Development, Washington.

Resignations

Harry Allan, University of Massachusetts: Syracuse University, Aug. 31, 1970.

Miloslave Bernasek, Boston University: Macquarie University, New South Wales, Australia, Sept. 1971.

Irwin Bernhardt, Pennsylvania State University: University of Waterloo, Ontario, June 1970.

Philip M. deMoss, University of Missouri, Kansas City: PMC College.

Muzaffer M. ErSalcuk, Purdue University, Sept. 1970.

Charles Gallagher, University of North Dakota: University of Calgary, Alberta, Aug. 9, 1970.

Lee C. Garrison, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Irving Gershenberg, University of Connecticut: Makerere, Uganda.

Peter J. Ginman, Boston University: State University of New York, Geneseo, Sept. 1971.

Eila Hanni, Colorado College.

Robinson Hollister, University of Wisconsin, Madison: Princeton University, June 1970.

David M. Johnson, University of North Dakota, June 1970.

Howard P. Kitt, Hofstra University, fall 1970.

Mark W. Leiserson, Yale University: International Labor Office, Geneva, Switzerland, Sept. 1, 1970.

Edward Lorentzen, Colorado College.

James B. McCollum, College of Industrial Management, Georgia Institute of Technology.

Harry M. Markowitz, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Howard Pack, Yale University: Swarthmore College, July 1, 1970.

Rein Peterson, Graduate School of Business, Columbia University, July 1970.

Michael T. Quinn, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Hans O. Schmitt, University of Wisconsin, Madison: International Bank for Reconstruction and Development, Washington, June 1970.

Philip J. Schreiner, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Tibor Scitovsky, Yale University: Stanford University, July 1, 1970.

Isidore Silver, University of Massachusetts: John Jay College of Criminal Justice, New York, Aug. 31, 1970.

David Singer, Hofstra University, fall 1970.

Ernst Stromsdorfer, Pennsylvania State University: University of Indiana, June 1970.

Ward Theilman, University of Massachusetts: Texaco Company, New York, Aug. 31, 1970.

Myron Uretsky, Graduate School of Business, Columbia University, June 30, 1970.

Peter Vaill, Graduate School of Business Administration, University of California, Los Angeles: University of Connecticut, July 1, 1970.

George C. Wang, University of Southern California: University of Tennessee, Aug. 1970.

Donald H. Woods, Graduate School of Business Administration, University of California, Los Angeles, July 1, 1970.

Miscellaneous

Edwin J. Holstein: president of the New York State Economics Association.



EMPLOYMENT SERVICES

NATIONAL REGISTRY FOR ECONOMISTS

The National Registry for Economists was established in January, 1966, to provide a centralized nationwide clearinghouse for economists on a year-round basis. It is located in the Chicago Professional Placement Office of the Illinois State Employment Service and is staffed by experienced placement personnel, operating under the guidance and direction of Regional and National Bureau of Employment Security Professional Placement officials, and in cooperation with the American Economic Association. It is a free service. There are no registration, referral, or placement fees. Application and order forms used in the Registry are available upon request from the: National Registry for Economists, Professional Placement Center, 208 South La Salle Street, Chicago, Illinois 60604.

AMERICAN ECONOMIC ASSOCIATION VACANCIES AND APPLICATIONS

The Association is glad to render service to applicants who wish to make known their availability for positions in the field of economics and to administrative officers of colleges and universities and others who are seeking to fill vacancies in the field of economics.

The officers of the Association take no responsibility for making a selection among the applicants or following up the results. The Secretary's office will merely afford a central point for forwarding inquiries, and the *Review* will publish in this section a brief description of vacancies announced and of applications submitted (with necessary editorial changes). Since the Association has no other way of knowing whether or not this section is performing a real service, the Secretary would appreciate receiving notification of appointments made as a result of these announcements. Those submitting such announcements have the option of publishing either name and address or a key number in the listing. Inquiries about a listing with a key number should refer to it specifically and be mailed to the Secretary's office. Resumes and application blanks are not supplied by the American Economic Association. The Association will only forward inquiries and resumes to the proper party for their consideration. Deadlines for the four issues of the *Review* are January 1, April 1, July 1, and October 1.

Communications should be addressed to: The Secretary, American Economic Association, 809 Oxford House, 1313 21st Avenue South, Nashville, Tennessee 37212.

Vacancies

Fishery economists: Wide variety of economics research, ranging from international agreements, quotas, and tariffs to price analyses, business management of firms, cost-benefit analyses, and whole field of the economics of natural resources. Positions are in the federal Civil Service at Grades GS-9 (\$9,320) to GS-14 (\$18,531). Basic requirements are Ph.D. or master's in economics or agricultural economics; training in international, natural resource, and/or quantitative economics would be helpful. Positions are located at the University of Maryland, in Washington, D.C., and field positions. Civil Service Commission Form 171 (Application for Federal Employment) should be sent to: Personnel Office,

Bureau of Commercial Fisheries, U.S. Department of the Interior, 18th and C Street, N.W., Washington, D.C., 20240.

Economist: Opening in the Department of State Planning. Master's degree in economics and six years of experience required. Additional graduate study may be substituted for two years of the required experience. Background in quantitative research methods desirable. Starting salary \$13,739 effective September 1, 1970, with a maximum of \$18,049 reached in six years. Write: Vladimir Wahbe, Secretary of State Planning, Department of State Planning, 301 West Preston Street, Baltimore, Maryland, 21201.

Business administration: Applications sought at all rank levels for rapidly expanding program, favorably located in Vancouver metropolitan area. Both undergraduate and graduate work, but emphasis on latter. Integrated with strong economics program; well established at B.A.-Ph.D. levels. Additional specialists sought in behavioral aspects of management, organization theory, finance, managerial economics, marketing, information systems, and business policy. Doctorate required. Write to: Dr. John P. Herzog, Acting Chairman, Department of Economics and Commerce, Simon Fraser University, Burnaby 2, British Columbia, Canada.

Senior appointment in economics: Newer university, attractively located in Vancouver metropolitan area, seeks candidates with notable academic record to supervise doctoral research programs and teach at various levels in substantial economics program with well-qualified students. Write to: Dr. John P. Herzog, Acting Chairman, Department of Economics and Commerce, Simon Fraser University, Burnaby 2, British Columbia, Canada.

Staff economist—systems analyst: The Center for Naval Analyses, operated under contract with the University of Rochester, has three-four openings for Ph.D. economists. CNA is engaged in a broad spectrum of research involving operations research and systems analysis for the U.S. Navy and government agencies. CNA analyses provide bases for decisions in the planning and budgeting of military forces and government operations. This involves the application of economic principles to resource allocation problems relating to the sizing, mix, procurement, and use of alternative systems. Responsibilities include defining and structuring problems of choice, identifying alternatives, analysis of alternatives and constraints, illuminating relative costs and benefits. We have recently studied: the naval ship procurement process, cost and effectiveness of land-based and sea-based tactical aviation, trade-offs between aircraft procurement and maintenance, and the supply of labor to the military in the absence of a draft. We now plan to study: the organization and operation of the Navy resource allocation process, the allocation of air space between Navy and other users, the design of strategic forces, the costs and benefits of military medical systems, and a worldwide review of the naval aviation shore establishment. Time is available for professional development and research directed toward improving and/or developing analytical techniques and decision processes. Starting salaries range from \$16,000 to \$24,000, depending on education and experience. Staff members are eligible for merit increases at least annually. Write to: Mr. Terry E. Harris, Professional Staffing, Center for Naval Analyses, 1401 Wilson Boulevard, Arlington, Virginia, 22209. CNA is an Equal Opportunity Employer.

Industrial relations: A Ph.D. in industrial relations or economics to join the faculty in the expanding graduate (M.A.) program in industrial relations at a private college in the East, starting in January, 1971. Rank and salary open. P365

Economics: The United Nations has a continuing need for economists for challenging research and operational posts at Headquarters, New York, in Geneva, Vienna, and in developing countries throughout the world. These posts cover the full field of economic specialization (econometrics, fiscal policy, foreign trade, finance, agriculture, mining, water resources, industrial management, development, urban planning, etc.), at all levels of responsibility and of varying duration, permanent or short term. Representative of the posts now under recruitment are the following. The gross salary for these positions ranges between \$12,380 and \$20,490 (\$9,666-\$15,118 tax free), exclusive of dependence and post allowances which may amount to an additional several thousand dollars. (1) ECE, Geneva: International economists with several years of research experience and particular qualifications in East-West European trade; knowledge of French or Russian required. (2) ECLA, Santiago: Industrial economist with technical-economic research experience and practical experience in Latin American industrial sectors; e.g., textiles and metal transforming; Spanish required. U.S. citizens interested in obtaining additional information about these as well as other UN employment possibilities are requested to send a résumé to: Office of International Organization Recruitment, Room 4336A (IO/IR), U.S. Department of State, Washington, D.C., 20520

Economist: Ph.D., economics, with research and publication experience. Position to encompass quantitative research in the Center for Business and Economic Research and teaching at the graduate or undergraduate level. Rank of associate or full professor with salary dependent upon qualifications. Please send résumé to: Professor J. F. Niss, Chairman, Department of Economics, Western Illinois University, Macomb, Illinois, 61455.

Economist, econometrician, air transport research director: Air transport industry association is establishing new position, Director of Economic Systems Research, to develop the quantitative and/or scientific approach to the many economic problems faced by scheduled air carriers. Advanced degree in economics and knowledge of econometrics are required. Location, Washington, D.C. Write to: Golightly & Company International, Inc., 1 Rockefeller Plaza, New York, N.Y., 10020.

Economist: The National Association of Supervisors of State Banks (NASSB) has a vacancy in the position of research associate to be responsible for continuing research program under direction of Executive Vice-President. NASSB is engaged

in a broad range of bank and bank supervisory research including bank structure and performance, bank code revision, examination procedures, and background data for congressional testimony and various litigation and maintenance of comparative statistical indices. Broad opportunity to initiate research projects. Prefers candidates who have recently completed Ph.D. or D.B.A. but will consider doctoral candidates (M.S.-M.B.A. minimum). Salary competitive. Write to: Dr. Harry P. Guenther, Executive Vice President-Economist, NASSB, 1101 17th Street, N.W., Washington, D.C., 20036.

Marketing, accounting and finance, engineering economy: Positions for associate professor of marketing and of accounting and finance and assistant professor of engineering economy. Ph.D. plus university and/or industrial experience. Teaching and research in areas of specialization. Salary competitive. Send résumé to: Chairman, Department of Engineering Management, University of Missouri, Rolla, Missouri, 65401.

Business administration: Rapidly growing branch of University of Alaska, offering B.B.A., B.A. in Economics, and M.B.A., requires D.B.A. or Ph.D. in business administration or dissertation stage. Nine-hour teaching load covering graduate and undergraduate programs plus supervision of graduate students. Rapidly expanding program in Anchorage, the largest population center in state. Research opportunities, summer teaching, plus chance to help formulate departmental programs and policies. Write to: Dr. R. Maurice Jones, M.B.A. Program, University of Alaska, South-central Regional Center, 1820 West Northern Lights Boulevard, Anchorage, Alaska, 99503.

Accounting: Two accounting positions open for September, 1971. Rank can be from instructor to professor. Salaries are competitive and will be commensurate with education and experience. Interested parties may write or phone at our expense. Contact: Dr. Lowell Chapman, Head, Department of Accounting, Ferris State College, Big Rapids, Michigan, 49307. Telephone: 616-796-8359, Ext. 526.

Economist: Opening for September, 1971, in small teaching-oriented liberal arts college. Prefers Ph.D. and some teaching experience. Fields needed include American economic history, micro, labor, money and banking. Competitive salary depending on qualifications. Generous fringe benefits and consulting opportunities in local "Model Cities" program. Please send résumé to: Dr. Dennis J. Mahar, Acting Chairman, Pikeville College, Pikeville, Kentucky, 41501.

Economics: Ph.D. or D.B.A. preferred to fill positions in recently established Department of Economics offering a major in economics to teach principles of economics and undergraduate "mathematical" economics and/or first-year ac-

counting. Rank and salary: assistant professor to professor, \$9,000 to \$16,000, depending on degrees and experience. Write to: Department of Economics, Edinboro State College, Edinboro, Pennsylvania, 16412.

Department chairman: A fifteen-man Department of Economics with undergraduate and graduate programs requires a man with Ph.D. and teaching and research experience as department chairman. Salary commensurate with academic qualifications and experience. Submit résumé to: J. W. Skinner, Dean, College of Social Science, University of Guelph, Guelph, Ontario, Canada.

Economics: Several young economists with doctorate completed for September, 1971. Various specialties but a strong theoretical or quantitative orientation expected. Knowledge of computer utilization desired of one new faculty member. Good teaching and some publication required for tenure and promotion. Favorable location in interior of northern California. Write to: John A. Orr, Chairman, Department of Economics, Chico State College, Chico, California, 95926.

Economists: Applications are invited for two or three appointments in September, 1971. Ph.D. required for professorial ranks. Please send résumé to: Dr. I. Ghurani, Chairman, Department of Economics, Pahlavi University, Shiraz, Iran.

Business administration: Applications are invited for one or two appointments in accounting and in general business. Ph.D. required for professorial ranks. Please send résumé to: Chairman, Department of Economics, Pahlavi University, Shiraz, Iran.

Finance and management: National Development Corporation, responsible for promoting development prospects in Tanzania, has vacancy for Group Loan Officer in Office of Planning and Finance. The Group Loan Officer ensures that the Corporation's loans are made in accordance with established policies and standards of financial prudence; manages the deployment of short-term funds between Group companies; and is responsible for providing continual guidance to NDC top management on prevailing conditions in worldwide money markets. Salary negotiable depending on qualifications. Write to: George Kahama, Managing Director, National Development Corporation, Development House, Azania Front, P.O. Box 2669, Dar es Salaam, Tanzania.

Economics, management science: Two new positions for 1971-72 with primary specialties in the areas of behavioral economics, finance, simulation models for decision making and economic theory. Appointments will be made at the assistant and associate level. Ph.D. or near Ph.D. required. Minimum salary A.A.U.P.-A. Address applications to: Dr. David A. Martin, Chairman, De-

partment of Economics, State University of New York, College at Geneseo, New York, 14454.

Marketing: Rapidly growing private, nonsectarian, coeducational college seeks chairman for expanding Marketing Department. Excellent potential for growth in pleasant work and residential environment on new suburban campus near Princeton, New Jersey. Applicant should have evidence of scholarly research and publication record. Training in behavioral sciences or quantitative methods and computer experience are desired. The applicant must have Ph.D. or D.B.A. with administrative and teaching experience. Teaching load includes undergraduate and graduate courses in areas of personal interest and training. Rank and salary competitive. Inquiries should be sent to: D. N. Dertouzos, Dean, School of Business Administration, Rider College, Trenton, New Jersey, 08602.

Business administration: The Business Administration Department of Mankato State College is seeking applicants for the position of department chairman. Applicant must possess an earned doctorate, teaching experience, and administrative supervisory or business experience is desirable. Salary \$20,000-\$22,500. Send application to: Dr. Morgan I. Thomas, Dean, School of Business, Mankato State College, Mankato, Minnesota, 56001.

Economist: Attractive openings in Washington, D.C., for individuals with Ph.D. or exceptional graduate school record without completed dissertation for position (two-three days per week) doing research on social welfare problems. Background in public finance and microeconomics preferred. Ideal for woman who wishes to combine domesticity and important work or faculty member who wishes to combine teaching with policy analysis. P369

Economics—general competence, interest in both undergraduate and graduate teaching emphasizing development economics. One or more specialties: theory, history, systems, labor, finance, money and banking, international trade, accounting, econometrics policy, computer application. Minimum four years post-doctoral experience English-medium institution, excellent teaching recommendations, professional and personal adaptability. Compensation comparable average U.S. institutions or willing sub-contract, two year appointments customary. Write Dean of Faculties, American University in Cairo, at NY Office, 866 UN Plaza, NY, NY 10017.

Economic Planning and Development—for undergraduate and graduate programs. Extensive and varied teaching and practical experience desired. General competence and additional specialties welcome.

Minimum four years post-doctoral experience English-medium institution, excellent teaching

recommendations, professional and personal adaptability. Compensation comparable average U.S. institutions or willing sub-contract, two year appointments customary. Write Dean of Faculties, American University in Cairo, at NY Office, 866 UN Plaza, NY, NY 10017.

Accounting: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics and management, has a position open at assistant or associate professor rank. Courses to be taught include principles and upper division electives. Creative teaching required. Teaching load 12 hours. Requirements: C.P.A. and M.B.A., and college teaching or other allied experience. Salary is commensurate with qualifications. For further details write to: Dr. Joseph M. Thorson, Chairman, Department of Business and Economics, School of Social and Behavioral Sciences, West Chester State College, West Chester, Pennsylvania 19380 or call 215-436-2236.

Business Administration: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics, and management, has a position open for an assistant or associate professor of management. Courses to be taught include principles and upper division electives. Creative teaching required. Teaching load 12 hours. Requirements: D.B.A. or Ph.D. in business administration and college teaching or allied experience. Salary is commensurate with qualifications. For further details write to: Dr. Joseph M. Thorson, Chairman, Department of Business and Economics, School of Social and Behavioral Sciences, West Chester State College, West Chester, Pennsylvania 19380 or call 215-436-2236.

Marketing: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics, and management, has a position open at the assistant or associate rank. Courses to be taught include principles and upper division electives. Creative teaching required. Teaching load 12 hours. Requirements: D.B.A. or Ph.D. in marketing and college teaching or allied experience. Salary is commensurate with qualifications. For further details write to: Dr. Joseph M. Thorson, Chairman, Department of Business and Economics, School of Social and Behavioral Sciences, West Chester State College, West Chester, Pennsylvania 19380 or call 215-436-2236.

Economics: Ontario University has several possible openings in various fields of economics. Duties include teaching and research in a quality four year undergraduate honours program. Ph.D. required. Rank and salary open. Send curriculum

vitae and names of three referees to: John A. Weir, Chairman, Dept. of Economics, Waterloo Lutheran University, Waterloo, Ontario, Canada.

Business Administration: Rapidly growing department of Central Washington State College has openings for 1971-72 in accounting; finance; insurance; real estate; quantitative methods; management; industrial relations; marketing; economics. Opportunity to participate in development of programs at bachelor and masters level. Ph.D., D.B.A. or dissertation stage preferred. Salary open and competitive. Small town location with easy access to metropolitan area. Outstanding outdoor recreation area. Write to Professor L. A. Danton, Chairman, Dept. of Economics and Bus. Admin., CWSC, Ellensburg, Wash. 98926.

Economist: Instructor or Assistant Prof. to teach principles of economics and two upper division courses at a New England L.A. College with 1800 students. The upper division courses are somewhat open but the department members are inclined towards adding courses in comparative systems, and/or urban and/or welfare. Ph.D. dissertation stage or Ph.D. is required. The teaching load is 12 hours with a competitive salary schedule. P370

The Graduate Program in Business Administration of the University of Alaska at Anchorage announces the following positions for the 1971-72 academic year. The new positions are in the areas of econometrics, macro-economics, resource-urban economics and finance. Positions are for 9 or 12 month contracts depending on the applicant. Applicants should possess terminal degrees in the areas listed above or be in the final stages of dissertation before seriously considered. Salary ranges for the Assistant through full Professor positions range from \$18,900 to \$28,500 for 12 month contracts. Applications to Dr. R. Maurice Jones, Head, Division of Business Administration, University of Alaska at Anchorage, 1820 W. Northern Lights Boulevard, Anchorage, Alaska 99503.

Economics: September, 1971. Ph.D. required. Rank: Assistant Professor. Salary: Open for negotiation depending upon teaching experience and background. Courses: Principles, Macroeconomics, Growth, and Elementary Econometrics. TIAA-CRBF, tax-sheltered program, social security, group life insurance, major medical, sabbatical leave, and faculty's children tuition benefit. Please send résumé to: Dr. C. F. Joseph Tom, Chairman, Department of Economics and Business Administration, Lebanon Valley College, Annville, Pennsylvania 17003.

Economic Education: Opening for qualified specialist in the area of economic education; primary responsibilities include the establishment and administration of a Center for Economic Education plus modest teaching duties; Ph.D. and some experience preferred; salary and rank open; liberal fringe benefits. Send resume to Dr. Eugene A.

Philipps, Chairman, Department of Economics, Moorhead State College, Moorhead, Minnesota 56560.

Dean, College of Business, Economics and Government: University of Alaska, Fairbanks. From \$27,000/year, according to qualifications. Doctor's degree; individual with proven academic experience and qualifications, with administrative abilities, and preferably with some research experience. To be responsible for administering the affairs of the College with a budget of approximately \$500,000 and approximately 50 professional faculty and staff. Contact Dr. Richard J. Solie, Chairman, Advisory Selection Committee for Deanship, College of Business, Economics and Government, University of Alaska, College, Alaska 99701. Telephone number: 907-479-7120.

Economics: Few positions open in a growing Economics Department. Ph.D. preferred; fields of specialization: urban economics, history of thought, economic development and planning. Salary open. College ideally located in rural setting close to eastern metropolitan centers. Openings are for fall term, 1971. Apply to: Dr. M. K. Hamid, Acting Dean, Division of Social Sciences, Millersville State College, Millersville, Pennsylvania 17551.

"Co-educational, liberal arts college in Central Pennsylvania is seeking to augment small economics department. Large percentage of majors pursue graduate work. Candidate must be strong in undergraduate teaching and desire to innovate, particularly in the principles course, with the hope of improving the teaching of political economy. Pedigree and publication secondary qualifications; teaching ability of primary concern. Rank and salary commensurate with qualifications." Write directly to: Robert W. Rabold, Chairman, Department of Economics, Lycoming College, Williamsport, Pennsylvania, 17701.

Chairman, Department of Economics: A small department of economics with growth potential is seeking a chairman with a PhD in economics, teaching experience, and some professional publications. Rank and salary are open for discussion. Please address inquiries to Dean Thomas Powell, College of Arts and Sciences, State University of New York, College at Oswego, Oswego, New York 13126.

Economists: The Department of Economics, The University of Calgary has vacancies for Senior Economists in the fields of Micro Theory, Econometrics, Petroleum Economics and Transportation. Present salary ranges are: Associate—\$14,630—\$19,665 with \$640 increments per year. The minimum salary for Professors is \$19,715. The University of Calgary has an enrollment of over 9,000 full-time students of whom ten per cent are graduates. The Department of Economics expects to obtain permission to offer the Ph.D. in 1971. Calgary has a population of 380,000, excellent cultural facilities and is located in the foothills of the

Rocky Mountains. Superb ski and camping facilities are readily accessible. Please send applications to: Dr. F. R. Anton, Head, Department of Economics, The University of Calgary, Calgary 44, Alberta, Canada.

Administration and Economics Department: 2 professors in Economics. Fields: the economic theory—research on the economic development of urban and regional areas. Qualifications: preferably Ph.D. Social benefits: life insurance, sickness insurance, salary insurance, pension funds. Salary: according to scale of salary presently in force at the University of Quebec at Trois-Rivières. Send curriculum vitae to: Mr. Andre Heroux, University of Quebec at Trois-Rivières, C.P. 500, Trois-Rivières, P.Q., Canada.

Presidency, Eastern Illinois University: Eastern is a state supported university offering work through the specialist with an enrollment of 8,600. Accredited by North Central, its 700 teaching faculty includes an unusually high percentage of earned doctorates. With an annual operating budget of approximately \$20,000,000, the university has a 316 acre campus and 37 major buildings. The person sought must have successful experience in all aspects of higher education and its administration. Address inquiries or nominations to: Board of Governors of State Colleges and Universities, 222 College Street, Springfield, Illinois 62706.

Quantitative Methods in Business; Marketing; Management; Finance: Ph.D. in Economics or Doctor of Business Administration (or candidates who are close to attaining one or the other of these degrees) sought for Fall, 1971, to teach in four-year-old, rapidly growing Master of Business Administration Program, as well as at the undergraduate level in Business Administration. Salary and fringe benefits fully competitive with other institutions; rank depends upon experience. For further information, make application to: Dr. W. P. Carton, Director, Master of Business Administration Program, Loyola College, 4501 North Charles St., Baltimore, Maryland, 21210; Telephone No.: 301-435-2500.

Econometrician: Ph.D. plus several years experience in creating and building mathematical models. Must possess knowledge of linear programming, simultaneous equations, matrix algebra and the IBM 360 Computer. Programming background helpful. Salary to \$20,677. Increased to \$22,691 April, 1971. Send resume in confidence to Personnel Officer, New York State Department of Public Service, 44 Holland Avenue, Albany, New York 12208.

Chief Statistician, Bureau of Business Research: September 1, 1971. Principal duties include (1) performance and supervision of research in economic and business areas, with concentration on population, income, and regional analysis; (2)

consultation and planning on statistical, econometric, and other quantitative matters in areas of economics and business. Some teaching if desired. Opportunity for research and writing in areas of personal interest. Research experience in areas described above highly desirable. Ph.D. desirable but not essential. Salary (11 months): \$15,000-\$16,000 maximum for 1971-72 plus substantial fringe benefits. Send application and resume to Dr. E. S. Wallace, Bureau of Business Research, 200 College of Business Administration, University of Nebraska, Lincoln, Nebraska 68508.

Economics: Openings for professor and associate professor in a rapidly expanding department with B.A. and M.A. programs in economics. Ph.D. degree, teaching experience at graduate level, and publication record required. Areas of specialization open. Teaching load favorable to research and writing. Salaries and fringe benefits excellent. Close to Rochester metropolitan area. Write to Dr. Joseph Mason, Chairman Economics Department, State University of New York, College at Brockport, Brockport, N.Y. 14420.

Labor economics, international trade, econometrics and mathematical economics. Will consider people at various ranks depending on their qualifications. Salary dependent upon qualifications. This position will be open July 1, 1971, for the academic year. Course work begins in mid-September, 1971. Contact: Dr. Maurice Perkins, Chairman, Department of Economics, Brock University, St. Catherine's, Ontario, Canada.

Statistics, monetary and fiscal policy. Openings for full and associate professors. Requirements are Ph.D., D.B.A., or M.B.A. with course work completed for doctorate and college teaching, business, or government experience. Publications desirable but not essential. Salary range depending on education and other qualifications; nine months with summer session optional; annual increments; high fringe benefits. A Catholic Liberal Arts College with a Division of Business, located in pleasant suburban surroundings in up-state New York. Send resume to: Chairman, Business Division, Siena College, Loudonville, New York 12211.

Economists: Expanding Economics Program has vacancies beginning September, 1971, in following areas: monetary theory and policy, economic history, comparative systems, econometrics, labor economics, public policy, regional economics and economic development. Ph.D. preferred with teaching competence, deep appreciation for economic theory and measurement, and a desire to help develop a fine economics program. Send full résumé to: Chairman, Economics Program, Southwest Minnesota State College, Marshall, Minn. 56258.

Economics: three openings for 1971-72 with main specialty in any of the following areas: 1) public

finance, 2) money and banking, 3) macroeconomics. Twelve hour teaching load that includes one graduate course (except for the Instructor position). Extra teaching loads with extra compensation are available during the academic year at extension centers with summer teaching or research a possibility. Positions are available at the following ranks: *Associate Professor*—Ph.D. in Economics *in hand* required with at least six years of college or university teaching experience and some publications. Salary range (nine months) \$11,500-\$14,000; *Assistant Professor*—Ph.D. in Economics *in hand* required with preferably some college or university teaching experience. Salary range (nine months) \$10,000-\$12,500; *Instructor*—M.A. in Economics plus work towards the Ph.D. in Economics and preferably some college or university teaching experience. Salary range (nine months) \$8,000-\$9,500. Send resume to: Dr. Joseph S. La Cascia, Chairman, Department of Economics, Marshall University, Huntington, West Virginia 25701.

Small liberal arts college in midwest, metropolitan area seeks an innovative economist with an interdisciplinary background to help develop a new approach to undergraduate economic education and education for the policy sciences. The position requires a dynamic individual committed to undergraduate education who currently holds the Ph.D. or is near it. Preferred, as areas of speciality, are urban and regional economics, natural resource economics, international trade, and labor. However, other areas will be considered. Compensation and rank are dependent on qualifications. Send resume to Don A. Ostasiewski, Thomas More College, Fort Mitchell, Ky. 41017.

Economics: Opening Sept 1971 in selective liberal arts college. PhD and successful undergraduate teaching experience. Flexible teaching assignment, most desirable areas are: general economics, business & government, industrial organization, and managerial accounting. R. K. Stuart, Chairman Economics Dept; Whitman College, Walla Walla, Wash. 99362.

Economics, Department Chairman: Large Private University in a metropolitan area is seeking an experienced economist with Ph.D. to head a ten man department. Publication record desirable. Rank and salary open and competitive. Send resume to: D. B. Stewart, Chairman, Personnel Committee, Department of Economics, University of Dayton, Dayton, Ohio 45409.

Business Administration/Economics: Three position at full or associate professor rank. Ph.D.'s or D.B.A.'s wanted. Growing School of Business will initiate an MBA program in 1971. Teaching assignment flexible in accounting, management, finance, and economics at both undergraduate and graduate levels. Associate salary \$11,000 to \$13,000; full, \$12,500 to \$15,000. Additional compensation includes subsidized housing and

travel. Send resume to: Dr. Edwin Carey, Dean, College of Administration and Technical Studies, University of Guam, Agana, Guam, 96910.

Research associate: A well-trained economist with a keen analytical mind is needed now for a permanent position to assist in projects dealing largely with public finance—federal and state taxation, budgets, major expenditure fields—fiscal policy, intergovernmental relations, employment, poverty war and related subjects. Master's degree and/or work experience in one or more of those subjects desirable but not essential. Starting salary \$10,000 to \$12,000 or more, depending on qualifications. Opportunity to teach, if desired. Write to: Roger A. Freeman, Hoover Institution, Stanford University, Stanford, California, 94305.

Economics: The "Institut für Weltwirtschaft" has several openings for qualified research economists, preferably specialized in development economics, who would like to participate in a study assessing the efficiency of industrialization policies in LDCs with a view to projecting their future export potential.

Starting salaries range from DM 20 000 to 30 000 depending on age, family size, and qualifications. Period of employment for at least one year, preferably longer. A basic knowledge of German would be of advantage.

Write to: Prof. Herbert Giersch, Institut für Weltwirtschaft, D-23 Kiel/West Germany, Düsternbrooker Weg 120.

Economist: September 1971. The Department of Economics invites applications for a position at the Assistant Professor level. Primary responsibility of the candidate will be to introduce a rigorous program in Quantitative Methods (Econometrics, Mathematical Statistics) into the major program for all Economics majors. Ph.D. and teaching experience preferred. Write to: Professor Lewis P. Fickett, Jr., Department of Economics and Political Science, Mary Washington College of the University of Virginia, Box 1036, College Station, Fredericksburg, Virginia 22401.

Economist: Expanding four year program requires a specialist in international trade and either economic development or comparative systems. Opening at the associate or professor rank with opportunity to develop program in economics and to participate in a new Division of International Studies. Rank and salary are open for negotiations. Write: Professor Thomas Smuck, University of Hawaii, Hilo College, Hilo, Hawaii 96720

Economics: Position open for the fall of 1971 in a small department requiring an interest in teaching introductory economics. Interest in business applications helpful for an anticipated program in business and economics. Near Ph.D. preferred. Appointment can be made at any level depending upon the candidate's qualifications and experience. Write to: Professor Thomas Smuck, Chairman,

Division of Social Sciences, University of Hawaii, Hilo College, Hilo, Hawaii 96720.

Staff Economist: An established, progressive research-planning organization seeks a practical working economist, Ph.D. preferred, to fill a newly-created position. Responsible to the firm's economic-social sector's growth and the preparation of economic-social reports within an over all variety of planning programs. Should be familiar with urbanization processes; social needs of communities and cities; ecological trends and problems; federal funding; private financing; planning on the local, city, regional, and state levels; building engineering; and related fields. Must possess an expertise in researching, compiling, and analyzing economic-social problems and recommending their resolutions on a practical level. Essential is the ability to work with a team of highly qualified specialists in supportive endeavors. Salary \$18,000 plus fringe benefits. Send resume to: Dr. William Bryan, President, Economic-Social Research Company, 975 East Hudson Street, Columbus, Ohio 43211

Economics: Two new positions in the area of marine resource problems and economic development in a Department with a strong graduate program in marine economics. Basically research positions, but opportunities for teaching as needs permit. Research support available and possibilities for international travel exist as research program develops. Positions are open at the rank of assistant professor with starting salary of \$14,000-16,000.

Economist Available for Positions

Italics indicate fields of specialization

Economics, Investments, Computer applications: Man, married, Ph.D. Over 15 years experience in investments, business, government, and teaching. Strong forecasting record, along with training in reporting, analysis, and decision making. Wish to develop computer applications. Experienced in both oral and written communication. High level position desired in investments or business economics. E1044

Principles—micro and macro analysis, consumer economics, business law, law and society, law and economics, labor economics, government and business, economic policy, public finance and fiscal policy, intermediate theory, comparative systems: Man, 42, married; A.B., A.M., LL.B., LL.M. Experience includes attorney-at-law since 1954; elective public office 4 years; 12 years college teaching, 6 as chairman of economics department; Ford Foundation Fellow in Economics, Wharton School; present rank associate professor. Member of Quakers, wife Japanese, require congenial environment, academic freedom; no teaching evenings, off-campus, or summers (must be free to travel, study, write); prefer small Christian liberal arts college, office help for mimeographing publications, \$15,000 salary. E1182

Candidates should have a Ph.D. in economics or agricultural economics and either have experience or dissertation in the area of Economic Development. Write to: Dr. A. Holmsen, Department of Resource Economics, University of Rhode Island Kingston, R. I. 02881.

Economist: Macroeconomics, economics of environment, or structure of industry. Some other specialization is a possibility. Ph.D. required. Please send resume to: James E. Jensen, Chairman, Department of Economics and Business, Wisconsin State University-Stevens Point, Wis. 54481

Marketing: Position open September 1, 1971. Associate professor or assistant professor. Full-time career position in teaching and research, with choice of courses in general marketing, advertising, consumer behavior, or marketing research. Development of new courses is encouraged in both quantitative and behavioral aspects. Work will be mostly with undergraduates, but there are increasing needs for graduate teaching and for guidance of masters' and doctors theses in a complete Graduate College. Salary competitive. Summer teaching optional for extra salary. Funds are available for research and related travel. College is member of AACSB and has computer facilities, short-course and institute programs. New offices and classrooms. Community is excellent for family living. Send personal data sheet, with references, to: Professor Phillip McVey, Department of Marketing, University of Nebraska, Lincoln, Nebraska 68508.

Economic growth, microeconomic theory, economic development, mathematical economics, international trade. Man, 28, married, publications, expect Ph.D. fall 1970, four years teaching experience, desires teaching position with some graduate-school contact. E1828

Economic and Industrial Development; Macro and Micro-Economic Development Planning (including feasibility studies and evaluation of projects); International Trade, Banking, Finance and Investment; Business and Economic Forecasting, and Economic Research and Analysis of Problems of Business and Industry at Home and Abroad: Man, 62; Ph.D., Economics and Statistics. Served as Economic Advisor in AID Missions in Jordan, Brazil, Guatemala, Libya and Pakistan, and as Industry Economist in U.S. Government Agencies in Washington, D.C. Served as Professor of Economics and Statistics in a well-known University in the West and directed business and economic research program of the University. Interested in Consulting Work or University Teaching. E1850

Economics, micro and macro finance management, general accounting: Man, 61, married, Ph.D.

Twenty years college and university teaching; 9 years administration as department head. Business experience. Various publications. Desires teaching, research, administration, singly or in some combination. E1888

Theory, statistics and econometrics, development, industry studies, international economics: Man, 37; M.A., M.B.A. (cum laude), Ph.D. expected March 1971 from the New School for Social Research; all requirements except dissertation completed. N.S.F. dissertation research grant received. Ten years of teaching experience. Publications. Consultant to United Nations Center for Industrial Development. Available June 1971. E1896

Economic principles, international economics, intermediate theory, finance, economic development, Latin American economics: Man, 47, married, U.S. citizen, Ph.D. Wisconsin. Eight years teaching experience at state universities. Nine years as economist with international organizations. Presently consulting and teaching extension in Washington, D.C. Journal articles. Seeks full time teaching position. Will relocate. E2028

Management science, capital budgeting, accounting micro-economics: Male, 43, married; B.S. (engineering), C.P.A., M.S., Ph.D. (business and economics), dissertation in capital budgeting. Twelve years of part-time and some full-time teaching of quantitative economic and management science courses with 23 years of industrial and business experience. Recent 6 years as head of operations research. Desires teaching and/or administrative position in industrial management science function. Prefer midwest or western location. E2043

Business Administration: finance, management science, mathematics, planning: Ph.D. in business administration, other degrees in science and economics. Extensive diversified teaching; business experience as business manager, purchasing agent, and consultant; currently professor and chairman of business division. Desire teaching and/or administrative position at undergraduate or graduate level; also interested in long range corporate planning and management development. E2080

Industrial Organization, Government Regulation of Business, Economic Regulation of Public Utilities and Transportation, Microeconomic Theory: Man, 28, Married, B.A. (Physics), Ph.D. (Economics), Phi Beta Kappa. Three years teaching experience in state university, one year research experience in Federal regulatory agency, Brookings Institution Economic Policy Fellow, other fellowships, several publications. Desires teaching position with good research opportunities in university or college in major metropolitan area on east or west coast. Available, September 1971. E2101

Econometrics, statistics, economic analysis, international economics, economic development of

Communist China: Man, 34, married, Chinese citizen; B.A., M.A. (German equivalent), Ph.D. (German equivalent). Five years of teaching experience and researching works. Presently teaching at a British university in the Far East; some publications (books and articles). Desires research or teaching position in the U.S. or Canada. E2149

Econometrics, statistics, mathematical economics, labor economics, micro- and macrotheory: Man, 39; M.A. (economics) and 46 quarter hours of Ph.D. credit from U.S. universities; M.A. (economic statistics), Delhi School of Economics, India; economic planning study in Netherlands. Eight years of university teaching plus seven years of research experience in government. Desires teaching and/or research position. E2150

Statistics, management, marketing, finance: Man, single; J.D., Ph.D. Many years of experience in teaching and research; author of fifteen books and numerous articles; currently full professor. Desires teaching and/or research position. E2151

Economics of underdeveloped areas, market structure and industrial organization, agricultural economics, macro- and microeconomics, international trade, history of economic thought, public finance, comparative economic systems, statistics: Man, 39, married; B.S., LL.B. (India), M.A. and Ph.D. from U.S. state universities. Nine years of college and research experience; recipient of fellowship; member of honor society in economics. At present professor of economics in a U.S. state college; wishes change for location improvement. Desires teaching and research at a college or university. Prefers the Eastern Seaboard, West Coast, or any area close to metropolitan area but will consider other offers. Will consider positions in international organizations. Résumé furnished upon request. E2152

Industrial marketing, marketing research, production management: Man, 45; degrees in science and business, Ph.D. Thirteen years of college teaching; eighteen years of industrial and consulting experience with metalworking and electronic industries. Heavy research and publication record. Full professor of business. Desires a university, in a metropolitan area, offering a combination of research and teaching. Résumé furnished upon request. E2153

Economics, finance, management, marketing: Man, 62; Ph.D. Teaching and business experience. Publications and research. Available immediately. E2155

International labor, economic development, macro- and microeconomics, labor relations, law: Man, 40; undergraduate degree in diplomacy (foreign service), law degree, M.A. (economics-international affairs), doctorate candidate. For-

mer legislative representative with major labor union; over ten years of experience with nonprofits and U.S. government in economic analysis. Desires position in these and related disciplines. E2156

Public finance, international economics, history of economic thought, economic history, general economics: Man, 36, married; Ph.D. (1961). Some publications; consulting activity with regional government's council in public finance. Seeks senior position or possibly chairmanship in publicly-supported institution in Midwest or West Coast. E2157

Labor economics, principles, labor law, collective bargaining, personnel and industrial management, purchasing, organization planning: Man, 51, married; B.A. (economics), *cum laude*, M.S. (business administration), doctoral candidate. Ex-field grade military officer; eighteen years of industrial relations and purchasing executive. Seeks teaching and/or administrative position in university located in U.S. or overseas. Available in summer or fall, 1971. E2160

Public finance, micro- and macroeconomics, economic policy, quantitative economics: Woman, 27; Ph.D., University of Hamburg, West Germany, 1970. One year as Fulbright student at University of Michigan. Desires college or university teaching or research position for next academic year in Sacramento area, beginning in September. E2161

Econometrics, mathematical economics, operational research, statistics: Man, 35, married, Indian; M.Sc. from India, M.Sc. from England. About ten years of experience of research as a statistician; presently in England. Seeks research, analyst, or teaching position anywhere in U.S. E2163

Economics, development, marketing, business law, labor economics: Man, single; B.Com., LL.B., M.Sc. (economics), LL.D., *cum laude*, D.D. Twenty-six years of teaching and research experience. Several publications. At present professor of business administration at an Illinois university. Desires teaching position. E2165

Macro- and microeconomics, economic development, welfare, development of education, health, housing, public finance, management: Man, 48, married; M.S. and submitted doctor's thesis. Specializing in economy of eastern Europe. Experiences include ten years of management of economic planning commission, followed by ten years of extensive research and economic and financial analysis of national development and state budget. Experience overseas. Publications. Willing to perform research and advising duties. E2167

Marketing, finance and accountancy, international business, economic principles, economic problems

of developing countries: Man, 37, married, Indian; Ph.D. (business finance), London School of Economics, Graduate Diploma in Business Administration, London School of Economics, M.Com., LL.B., B.Comm., Calcutta. Eight years of teaching at universities in Africa. Business experience in India. Publications include a book and articles in journals. Seeks any suitable teaching and research position in U.S. or Canada. Presently serving in Africa; prepared to pay own removal and other expenses. E2173

Management consulting, accounting, finance: Man, 42, married; M.A. (economics), Ph.D. (business administration). Experience includes six years of teaching at graduate school level; ten years of consulting in accounting, finance, and corporate planning with extensive exposure to data processing. Currently employed as consultant with business. Seeks position in business or research offering creative challenge or administrative position with university. Currently located on West Coast. E2174

Economic development, marketing, international trade and finance: Man, 45, married; thesis pending for master's degree, economic degree from University of El Salvador. Experience includes over twelve years in consumer, industrial, and agricultural goods. Bilingual Spanish-English. Desires position in economic development, international trade. Will relocate in the U.S. E2176

Economic development, international economics, principles, history of economic thought: Man, 29, married; Ph.D. Four years of teaching experience; four years of government service; four years in Latin America. Wishes to leave foreign service for teaching/research position, preferably in Midwest university with Latin American studies program. Publications; fluent Spanish. Résumé on request. Available June or September, 1971. E2178

Monetary and fiscal theory and policy, finance, investments, marketing, forecasting, comparative economic systems: Man; Ph.D., Columbia University. Teaching and research since 1956. Last eight years with federal government in Washington, D.C. Publications. Fluent in five languages. Desires teaching position. E2180

Securities markets, investments, corporation finance, insurance: Man, 55, married; M.A. Fifteen years of teaching experience and a like number in industry. Presently associate professor of finance but desires to make a geographic change to a more challenging and specialized position. E2181

Mathematical economics, econometrics, statistics, input-output economics, microeconomic theory: Man, 34; Ph.D. Several years of full-time teaching experience. Currently assistant professor at a state university. Prefers teaching in university

with good graduate program but would consider any challenging research position. Available in September, 1971. E2182

Principles of economics, money and banking, statistics, public finance, international economics, economic development, labor economics, economic thought: Man, 30, Indian; M.A. (1st class), near Ph.D. Eight years of college and university teaching experience; presently professor of economics in India. Research experience; various publications, including research papers and books. Member of learned societies. Seeks teaching and/or research position anywhere in the world. E2183

Public finance, money and banking, price theory, corporation finance: Man, 50, married; Ph.D. Fifteen years of college teaching and ten years as research associate in public finance in large eastern university. Seeks teaching position with research opportunities or research responsibilities. E2186

Business finance, microeconomics, financial management, accounting: Man, 33, Indian; Ph.D. in business finance from a leading Indian Institute. Currently working as a lecturer in finance teaching M.B.A. classes in a premier university in England. Previous teaching experience in financial accounting and management accounting for five years in India. Coming to U.S. on an immigration visa in September, 1971. Seeks a faculty position in U.S. in September, 1971. E2187

Economic development, comparative economic systems, labor economics, collective bargaining, industrial relations, personnel management, manpower and human resources development, management, organization theory: Man, 48, married, Indian permanent U.S. resident; M.A., Ph.D., M.P.A., Harvard and M.I.T. Twenty years of university teaching, professor-director, graduate school of labor management studies with a big faculty; held senior civil service positions in Ministry of Labor and Social Welfare and as economic adviser in developing countries; worked as a director of management consultancy, training and research institute. Taught in many Asian, African, American universities; sixteen books and other publications. Currently teaching as a full professor at a reputed big midwestern university, graduate and undergraduate courses. Desires a university teaching position with research or administration or research position in government or other agencies. E2188

Economic development, econometrics, comparative economic systems, international trade, theory: Man, 28, married; B.A., M.A., Ph.D. expected early 1971. Former Peace Corps Volunteer (Chile), interested in Latin America. Fluent in Spanish; some Portuguese. Three years of teaching experience. Prefers teaching position in western U.S. Available in September, 1971. William Loehr, Department of Economics, University of Colorado, Boulder, Colorado, 80302.

Economic development, industrial organization, technological change, international law: Man, 32, married; B.S.E., LL.B., M.A., M.Phil., Ph.D. expected in 1971. Two years of experience with U.S. Agency for International Development in South Asia. Seeks overseas consulting, research, industrial, or governmental position where range of education and experience can be applied. Résumé on request. E2189

Regional economics, microeconomics, marketing, input-output analysis, management science: Man, 28, married; B.A., M.B.A., Ph.D. (business and economics). Six years of teaching and research experience; some publications. Desires teaching position in college or university. Vita and references upon request. Presently U. S. Army officer. Available in September, 1971. E2190

Micro- and macroeconomics, principles, mathematical economics, comparative systems: Man, 23; B.S. (honors), M.A. December, 1970, University of California, Riverside. Strong desire to teach and work for Ph.D. after gaining some experience. Available in January, 1971. E2192

Statistics, principles, international economics, economic development, mathematics for economists, econometrics: Man, 33, married; M.A., Ph.D. candidate, dissertation in progress. Six years of university teaching experience. Desires teaching position. Résumé furnished upon request. Available in June or September, 1971. E2194

Comparative economic systems, economic history, history of capitalism and socialism, micro- and macrotheory, economics of education: Man; Ph.D. Phi Beta Kappa; Fulbright Scholar. Thirteen years of teaching at graduate and undergraduate levels; publications; seven years of federal service (Bureau of the Budget, Foreign Service, State Department). Presently at state college. Prefers post in more progressive institution with A/S graduate program and larger library. Listed in *American Men of Science, Who's Who in American Education, Who's Who in the East.* E2195

General economic theory, monetary and financial theory and institutions: Man, 29, married, Canadian; Ph.D. Four years of teaching experience. Desires graduate and undergraduate teaching and research position. E2196

Macro- and microeconomics, mathematical economics, history of economic analysis, economic development and planning: Man, 43, married, Indian; State Doctorate from the University of Paris. Many years of teaching experience; presently teaching at a leading All India Institute. Fluent in French and knowing two other continental languages. Publications in standard journals. Desires teaching position in English- or French-speaking university. E2197

Econometrics, mathematical economics, operations research theory, money: Man, 31, married; B.A. (applied mathematics), Ph.D. (economics). Six

years of college teaching; three years in operations research; three years as an economist with a U.S. government agency. Currently a visiting professor in a major European country under a Ford Foundation grant. Desires research, teaching, or administrative appointment. Available in July, 1971. E2198

Principles of economics, micro- and macrotheory, monetary economics, labor economics, labor relations, finance: Man, 31, married; B.B.A., M.A., course work complete for Ph.D. Six years of teaching experience at large state university. Desires teaching position at four-year college where emphasis is on strong teaching. School near to medium-sized or large metropolitan area preferable but not essential. Would willingly accept varied teaching assignment. Good references. Available in fall, 1971. E2199

Finance, urban and regional economics, microtheory, operations research, mathematical economics: Man, 36, married; U.S. citizen; M.S. Chem. Engg., M.A., Ph.D. (economics). Over twenty-five publications, including coauthorship of three books. Presently professor and chairman at a major midwestern university. Eight years of teaching experience plus several years of administrative experience. Interested in mathematical programming and its applications to the economics of plant investment, capital budgeting, plant location, urban and regional economics and transportation. Seeks faculty appointment with or without administrative responsibility. Other types of positions also considered. E2201

Micro- and macroeconomics, economic development, monetary economics, history of economic thought, public finance, economic history: Man, 26, single, permanent U.S. resident; M.A., diploma in social sciences research methodology. Broad experience in research and sales. Presently employed as a sales representative in New York telephones. Interested in progressive research or sales positions. Can contact: Mohd. Anis, 2074 Harrison Avenue, Apt. 3-D, Bronx, New York, 10453.

Health economics: Man, married; recent Ph.D., large eastern university, honors. Publications. Broad teaching experience in economics. Interested in research, preferably but not necessarily with some teaching responsibilities. E2202

Theory, statistics and econometrics, international economics, industrial development and planning: Man, 36; M.B.A., M.A., Ph.D. expected February, 1971. Publications; consultant to international agency; ten years of teaching experience. Seeks teaching or research position beginning summer or fall, 1971. E2203

Economic theory, mathematical economics, statistics, econometrics, agricultural economics, operations research, economics development, public

finance, money and banking: Man, 37, married; M.S., M.A., Ph.D., University of California, Berkeley. Five years of full-time and one and a half years of part-time teaching in college and university to graduate students. Currently an associate professor. Minor publications. Desires teaching and/or research position. Available on reasonable notice. Résumé on request. E2204

International economics, economic systems, economic doctrines, economic planning and development, economic and socioeconomic and strategic analyses of foreign countries and areas, international and regional studies: Man, married; Ph.D. Worldwide diversified experience, including teaching. Desires administrative and/or teaching position developing integrated international and area studies forming responsible cadres for international careers, within a renowned institution of higher education or large corporation or organization with substantial international interests. Available in June, 1971. E2205

Economic development and planning, macro- and microeconomics, international economics, econometrics, industrial organization: Man, married, U.S. permanent resident; M.S. (mathematics), M.S. (statistics), Ph.D. (economics) from an outstanding American university, 1970. Fellowships, awards; several publications. Three years of teaching graduates and undergraduates in the U.S. and eight years of research abroad in a reputed academic institution and government. Desires relocation. Available after May, 1971. E2206

Macro- and microeconomic theories, international economics, money and banking, business cycles, public finance, quantitative economics: Man; Ph.D. Nine years of college teaching; Ford Foundation Research Fellowship; modest publication record. Presently associate professor in a state college. Desires a change in geographical location. E2207

Econometrics, mathematical economics, microeconomics: Man, 39; Ph.D. Ten years of experience in university teaching and research, including dissertation supervision; also experience with government and private research agencies; publications in applied area; interests in both applied and theoretical econometrics. Desires teaching and research or full-time research position but would consider other interesting positions. Available in summer, 1971. E2208

Simulation, analysis of dynamic systems, modern control theory: Man, 43; Ph.D. (engineering), U.C.L.A., 1967. Twenty-one years of experience in analysis, synthesis, and simulation of dynamic systems. Interested in variable and parameter estimation, optimization. Desires new career involvement in large socioeconomic system modeling and research or postdoctoral engineering and/or economics appointment with opportunity to advance in econometrics. E2209

Econometrics, mathematical statistics, computer-oriented applied economics: Man, 35, married, German national; doctor of economics (dissertation in econometrics). Ten years of academic research experience in econometrics; publications and several Econometric Society papers. Three years of teaching experience at major German university. Present rank roughly equivalent to associate professor. Desires a challenging affiliation with a U.S. institution (not necessarily a university). E2210

Economic theory, industrial organization, government and business, comparative economic systems, social ethics: Man, 37, married; M.A. (ethics), M.A. and additional graduate work in economics from a leading midwestern university. Eleven years of teaching experience in economics and introductory social science. Desires teaching position at a liberal arts college, preferably in the East or Midwest. Available in June, 1971. E2211

Public finance and fiscal policy, statistics, introduction to computer programming, macroeconomics: Man, 35, married; Ph.D. Five years of teaching experience and one year of experience as a government economist. Desires teaching position in a metropolitan area. Available in September, 1971. E2212

Economic and business history, comparative systems, growth and development, economic doctrines, public finance: Man, 33, single, widely travelled, U.S. permanent resident; M.A., Ph.D. from leading German universities; nine years of research and graduate teaching experience, last three years at outstanding West Coast campus; book, articles, reviews; excellent references. Desires university teaching and research position in U.S. or Canada; will attend 1970 A.E.A. convention; available in fall, 1971. E2213

Principles of economics, macro- and microeconomics, statistical methods, international economics: man, 26, married; B.S. (industrial administration), M.A. (economics), thirty hours toward the Ph.D. National Defense Fellow. Two years teaching experience. Desires college or junior college teaching position anywhere in U.S. Available January or June 1971. Credentials sent upon request. E2214

Money and banking, international trade and finance, development economics, economic thought, labor economics, principles: Man, 58, married; Ph.D. Wisconsin; 23 years of college teaching; chairman of economics department; widely travelled in Europe; linguistic ability. Seeks responsible position in a liberal arts college, preferably in the Midwest. Would also consider attractive offer for overseas position in teaching or with business firm. Available in September 1971. E2215

"Econometrics, Mathematical Economics, Economic Theory, and Government Regulation of Business. Man, 38, Harvard Ph.D., 1962. Extensive teaching and research experience. Interested in a teaching position with load not exceeding two courses per semester, or senior research position in Econometrics." E2216

Economics of education and health, public finance and policy, micro- and macroeconomic theory, principles, economic statistics: Man, 29, married; B.A., M.A., Ph.D. Experience in university teaching and research; several publications; substantial research in progress in the economics of education and health. Wish to relocate, to a university offering graduate work in economics. Location in Midwest preferable. Available July 1971. E2217

Economic theory, money and banking, government and business, management, finance: Man, 57, A.B., The College, University of Pennsylvania; A.M. Graduate School of Arts & Sciences, University of Pennsylvania, Philadelphia; Ph.D. Graduate School of Business Administration, New York University. Extensive college teaching and administrative experience as Department Chairman and Director of a School of Business Administration. Some modest publications. Presently teaching in Midwest but would like to relocate in South. Prefer Chairmanship of Department where teaching duties expected but am interested in any challenging building job. Resume on request. E2218

Development; money and banking, Southeast Asia studies: Man, 27; Ph.D., Cornell. Canadian with immigrant visa to U.S. anticipated early in 1971. 3 years teaching experience; some publication. Seeks university teaching position anywhere or research position with an international agency. E2219

Natural resource economics, regional economics, micro- and macroeconomic theory, managerial economics, monetary economics comparative economic systems, principles: Man, 34, married; Ph.D. Several years of full and part time teaching and research experience. Currently assistant professor at a small, state supported, liberal arts college. Desires teaching and/or research position with emphasis in natural resource and environmental economics. Available May, 1971. E2220

Economic Theory, Econometrics, Statistics, Monetary Theory and Policy, History of Economic Thought: Man, 26, Single; B.S. cum laude, Chemical Engineering; Ph.D. Economics to be obtained by Summer, 1971 (dissertation in progress). Scholarships, fellowships, and many other scholastic honors. Bilingual Spanish-English. Four years of teaching and research experience. Desires teaching position in U.S. Available in fall, 1971. E2221

Public Finance and Fiscal Policy, Micro and Macro Theory, Economic Development and Planning, Agricultural Economics and International Agriculture: Man, 35, Married, Indian, M.A. (Bombay), LL.B., M.A. (U.S.A.), Ph.D. dissertation in progress and degree expected soon. More than eight years of teaching and research experience; publications includes articles and a book. Desires teaching and/or research positions in the U.S. or Canada or International organizations. Available in June 1971. Resume furnished upon request. E2222

Economic theory—Micro and Welfare, Growth and Development. Man, 27, married, Ph.D. dissertation in progress (Vanderbilt); currently on Ford fellowship conducting research on pricing and investment in electricity industry in Latin American country; languages, will consider positions in U.S. or abroad. Available in June, 1971. E2223

Labor economics, collective bargaining, micro-economics, macroeconomics: Man, 24, married; M.B.A., Ph.D. One year full time and 2 years part time teaching experience, N.D.E.A. Fellow, desires teaching position in U.S. or Europe. Available in fall, 1971. E2224

Principles, labor economics, industrial organization, public finance: Man, 37, married; Ph.D. Ten years of teaching experience including five years as department chairman. Desires teaching position. Would consider chairmanship. Available in summer or fall, 1971. E2225

Public finance, tax policy and institutions, fiscal policy, allocation of resources in government sector, performance budget techniques, public investment project analysis, project management. Man, 39, Ph.D., Turk, fourteen years of university teaching (four years in an American College), adviser to the government, and research fellow to the UN. Resumé on request. Seeks teaching and/or research position in a university of English or French language. Available in May 1971. E2227

Public Finance, micro- and macroeconomics: Man, 35, Ph.D. Considerable teaching and research experience, some publications. Desires teaching position with opportunity and facilities for research in the area of public finance, particularly state and local finance. Available July, 1971. E2228

Money and banking, accounting, public and corporation finance, land economics, comparative economic systems: Man, 42, married, Canadian; B.A., M.A., some work towards Ph.D. Seventeen years of diversified experience in North America and overseas as an economist and accountant in industry, government, and a major university. Research, teaching, and other work. Fluent in German. Seeks permanent position on West Coast of

U.S. or Canada with government, bank, university, or other organization. Presently employed in Canada; visa for immigration to the U.S. assured immediately, if needed. E2229

Micro, regional economics, agricultural economics, human resources, planning: Man, 35, married. B.S., M.S., Ph.D. Experience includes two years of government research, five years of university teaching, research, and thesis supervision at assistant and associate professor levels, and consulting in planning and economic development. Publications and honors. Present position and salary satisfactory but desire relocation outside southeast and administrative responsibility. Prefer chairmanship of small or medium size economics department emphasizing quality teaching. E2230

Economics of American industry, micro- and macroeconomics, comparative economics, Soviet-type economics, economic thought, economic history of Europe, economics of natural resources, international trade. Man, 56; Ph.D. (Economics), M.A., M.C.L. LL.D. Experience in teaching industrial organizations, extensive research, international trade, lectures. French, German and Slavic Languages. Publications. Resume and List of Publications available. Prefers teaching and research; will consider government or industry. E2231

Micro and macroeconomic theory, economic development and planning, international economics, comparative economic systems, economic policy, economic integration (thesis and dissertation in this field): Man, 30, married, Jordanian; B.Sc. (Hons.), Cairo, A.M. and Ph.D., University of Southern California. Excellent record in undergraduate and graduate study. Three years teaching in the U.S. and Jordan. Two years experience in economic research at the Central Bank of Jordan. Modest publications. Desires a teaching and/or a research position. Now available. Tayseer A. Jaber, Central Bank of Jordan, P.O. Box 37, Amman, Jordan. E2232

Industrial relations, labor economics, international labor problems, economic history. Man, 31, Ph.D. from leading university. Teaching experience, publications, research experience in Latin America. Now employed in policy position with international organization. Desires teaching position in U.S. or Canada, fall 1971. E2233

Man; 35; married; Ph.D.; six years in Federal Reserve System and major commercial bank; experience in international economics, finance, development; publications; some capability in Spanish, French; references; seeks position in private-sector economics research or teaching post; prefers Washington, D.C. or central California. Will consider other locations; available early 1971. E2234

Public finance, urban and regional economics, human resources economics, economic development: Man, 29, Married, A.B., A.M., Ph.D. in progress. Government research experience; some publications. Desires teaching or research position. Available June or September, 1971. E2235

Economic development, international trade and finance, comparative economic systems, micro and macrotheory: Man, 27. Ph.D. expected June 1971. Two years teaching experience, two years research background. Fluent Spanish. Desires teaching position. Available September, 1971. E2236

Principles, macro- and microeconomic theory, money and banking. Man, 25, married; B.A., M.A. to begin Ph.D. course work in one year. Seeks research or teaching position in Michigan. E2237

Economic theory, economic history, history of economic thought, industrial organization, business management, labor economics: Man, married, B.S., M.S. (Bus. Admin.), M.A. (Economics), P.Eng., Ph.D. 1970. Industry experience as electrical engineer, contract administrator, manager, negotiator of government contracts: taught economic theory at University of Southern California part time. Seeks administrative and/or teaching position. Will relocate. Resume on request. E2238

Monetary theory, macro and micro theory, history: Man, 34; Ph.D. Currently at Ivy League university. Experience in teaching and research. Publications. Desires position as department chairman or with opportunity to teach in an active Ph.D. program. Preference for warm climate. Available June or September, 1971. E2239

Micro and macro theory, mathematical economics. Man, 40, Ph.D. from major university. Good publication record. Substantial graduate and undergraduate teaching experience. Available September, 1971. E2240

Economic development, introductory and intermediate micro and macro theory and policy, psychological economics, international trade, social science foundation of economics. Man, 40, married, B.S., B.F.T., M.B.A., Ph.D. expected March 1971. Six years business experience as market analyst, planner and economist with international divisions of two large U.S. firms; seven years college teaching experience in two and four year institutions; three years as economics consultant to major U.S. publisher. Experience in the development of instructional materials for introductory macro-micro sequence, especially television. Two years as newspaper columnist. Veteran desiring position in medium size academic institution which seeks to innovate at the under-

graduate level especially with regard to computer assisted instruction. Interested in teaching at undergraduate level with research opportunities, preferably regarding socio-economic development in Asia, Africa and Latin America. Resume on request. Available September 1971. E2241

Specialist in monetary theory, money and banking, 32, Ph.D. (econ.). Assoc. Prof. with 5 years full-time experience in teaching; 3 years full-time experience with Federal Reserve; consulting and research with investment services, Federal Reserve and university research center. Recent publications in *Quarterly Journal of Economics*, *Journal of Finance*, *Land Economics*, and *Journal of Financial and Quantitative Analysis*, among others. Resume on request. Available September, 1971. Bruce C. Cohen, 134 Lexington Ave., Cambridge, Mass. 02138. E2242

Economist Available Corporation Finance, Economics, Investments: Man, single; Ph.D., University teaching and research experience. Desire teaching position. E1765

Environmental, Urban, Labor Economics; Economic Development: Man, 48, married, MA in Economics. Experience includes 2 years as staff economist on multi-disciplinary environmental design concept team; 3 years in urban economics and economic development; and 13 years as labor economist and manpower specialist. Some teaching experience. Desires Federal Government employment. Eligible for GS-15 rating. Location in Washington, D.C. area only. E1844

Economics, micro and macro, finance, management, general accounting. Man, 61, married, Ph.D. Twenty years college and university teaching; 9 years administration as department head. Business experience. Various publications. Desires teaching, research, administration, singly or in some combination. E1888

Urban economics, U.S. and foreign economic development, housing market: Man; A.B., M.S. Forty years of high-level practical experience; widely traveled; knowledge of Spanish and Italian. Current security clearance. Fully-equipped private office in Washington, D.C. Available for a limited number of additional assignments on a contract, retainer, or per diem basis. Brochure on request. E1994

Economist: Man, 27; Publications; B.B.A., M.A., Ph.D. candidate with dissertation in progress. Area: Urban Economics. Background includes six years of extensive teaching experience in macro, micro, history of thought, statistics, urban economics and accounting. Seeks position on the college level as of September, 1971. E2132

Urban and regional economics and planning, input-output analysis, econometrics, statistics, economic development of Latin America: Man, 29, married, Ph.D. expected June, 1971, Cornell University. Presently teaching assistant in statistics and quantitative regional analysis. Two years of volunteer service and one year of dissertation research in Latin America. Desires teaching or research position in urban economics, planning, or economic development. Available after June, 1971. Resume on request. Robert A. Lewis, 526 Linn St., Ithaca, New York 14850 E2162

Mathematical Economics, Econometrics, Operational Research and Statistics: Indian, 36, married, M.Sc., from India, M.Sc., from England, ten years research experience as a Statistician, at present Lecturer at a British University, Green card holder—Permanent resident visa status, can take up assignment during summer 1971, seeks research, analyst or teaching position anywhere in U.S.A., prepared to consider visiting lectureship during summer or temporary position. E2163

Micro-macroeconomic, econometrics, statistics, labor economics: Man, 33, married, B.A., M.A., all Ph.D. requirements complete except dissertation. Six years undergraduate teaching experience which includes the initiation and supervision of the installation and operating policy of computer facilities to aid in the teaching of Business and Economics. Desires position at college with stress on teaching. Would prefer to locate on West Coast, Southwest or South. Available September 1971. E2179

Regional economics, economics of human resources, monetary economics, public finance, economic development, economic theory. Man, 32, Indian; presently working on Ph.D. dissertation in a leading Canadian university, also teaching at an affiliated college of the university. Several years of teaching and working experience in government. Desires teaching or research position in university or government or research organization. Available in July, 1971, or earlier if necessary. E2246

Planning Theory, urban and regional economics, resource economics, public expenditure theory, statistics, operations research: Man, 26, married, B.C.E. (Civil Engineering), Ph.D. dissertation in progress, awards and fellowships, research and experience in quantitative analysis of public policies and programs in above cited substantive areas, some publications. Desires teaching and/or research position. Resume on request. Available August, 1971. E2247

Population and developmental economics: Man, 46, married; A.M., Ph.D. (economics) Harvard, Certificate (demography) Princeton. Fourteen years of teaching experience, research in Latin

America. Presently associate professor. Seeks University position, temporary or permanent, to teach courses in demography, economics of population, and economic development. *Curriculum Vitae* on request. E2248

Planning, Evaluation and Systems Analysis: Man, 33, married; B.S., M.A., in quantitative economics. Seven years experience Federal Government in statistical analysis, model building, and program planning and evaluation. Interested in similar position in university administration, state or local government, or non-profit organization. E2249

Industrial Organization, Public Control of Business, Economic Theory, Economic History and Development: Age 34, B.A., M.A., Ph.D. in Economics. Three years teaching experience. Currently Assistant Professor at large state university. Administrative experience. Desires position as Chairman of small department or teaching position in Economics, preferably in west. E2250

Economic planning, transportation, labour economics, economics of public enterprise and agricultural economics: Man, 34, Indian migrating to the United States in February 1971, M.A. (Economics); six years' research experience in Indian Government's Planning Commission. Interested in joining International/American economic research organisation as Economist/Research Assistant. Available in New York, Feb. 1971. E2251

Economics, tax analysis, computer system design and applications, survey design, statistics: Man, 32, married; M.G.A. Wharton, Ph.D. Columbia. Recently studying business in urban ghetto. Consultant in programming in systems (O.S.) and applications, experienced in survey design, research for management, and teaching of undergraduate courses. Excellent references. Desires administrative or research position. E2252

Managerial economics, micro-economics, location theory, management, business policy: Man, 45, married; B.B.A., M.A., Ph.D. Five years business experience, sixteen years of University experience including positions primarily in teaching, research, and administration. Has taught on a team and in interdisciplinary courses. Prefers faculty appointment with or without administrative responsibility. Listed in *Who's Who in American Education*, *Who's Who in the Midwest*. Available June or September 1971. E2253

Economic development, econometrics specifically related to economic development, international trade and finance: man, 42 years, married, Ph.D. of famous European University with wide experience in academic work in Europe, U.S. (former associate professor Stanford Univ.) and Latin

America. After a five year field-work period in Latin America, desires to resettle in the academic life of the U.S.A. or Canada. Excellent recommendations, several publications, one book translated into two languages, other book pending.

E2254

Quantitative economic and business analysis, spatial economics, economic theory and planning: Man, 28, married; M.B.A., Ph.D. Three years teaching; currently associate professor of economics at a state university. Three years industrial operations research experience. Seeks teaching-research position or highly challenging consulting.

E2255

Money and Banking, Business Finance, Investments, Insurance, and Real Estate: Man, 62, mar-

ried; Ph.D. Fourteen years teaching experience at undergraduate and graduate levels. Publications. Business experience in finance field. Currently associate professor of finance. Desires to locate in New England or Middle Atlantic States. Available September 1971.

E2256

Monetary Economics, International Finance, Public Finance, History of Economic Thought, Economic Development, and Principles: Man, married, 36, A.M., Ph.D. (expected by June 71) both degrees from a prestigious eastern University. Four years of banking experience abroad, and seven years of college teaching in the U.S. Interested in a college or University teaching position. Available September 71. Excellent references.

E2257

New Penguins in ECONOMICS

PRINCIPLES OF DEVELOPMENT AID. *E.K. Hawkins.*

A guide to the study of outside aid for economic development—written for those who want to understand the subject as a part of the whole corpus of economic principles. \$1.65

THE CONTROL OF THE MONEY SUPPLY. *A.D. Bain.*

A complete survey of the methods, functions, and mechanism of monetary controls. \$1.95

JEVONS: THE THEORY OF POLITICAL ECONOMY. *Edited with an introduction by R.D. Collison Black.* \$1.65

THE MAGHREB IN THE MODERN WORLD. *Samir Amin.*

A look at the economic problems of Algeria, Tunisia, and Morocco from the late colonial period to the present day. The basic question: Is economic freedom possible without political independence? Available April. \$2.25

DEVELOPMENT IN A DIVIDED WORLD. *Edited by Dudley Seers.*

The problem of economic development in the world's poorer countries. Available May. \$2.45

SOVIET COMMUNISM AND AGRARIAN REVOLUTION. *Roy D. and Betty A. Laird.*

The authors contend that state and collective farming is grossly inefficient, and they go on to relate this failure to the deepening crisis in world food production. \$1.25

OIL AND WORLD POWER. *Peter R. Odell.*

The political and economic influence of the world's largest industry. \$1.25

INDUSTRY AND EMPIRE. *E.J. Hobsbawm.*

Volume 3 of The Pelican Economic History of Britain. Ranges from 1750 to the present. \$1.95

Penguin Modern Management Readings

... brings together outstanding articles and extracts, presented and edited by distinguished authorities, from every branch of management work and education. Each volume of Readings contains General and Part Introductions, a Further Reading list, and Author and Subject Indexes. Some recent titles:

MANAGEMENT AND MOTIVATION. *Edited by Victor H. Vroom and Edward L. Deci.* \$2.95

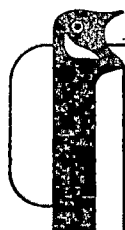
MANAGEMENT OF PRODUCTION. *Edited by M.K. Starr.* \$2.95

MANAGEMENT DECISION MAKING. *Edited by Lawrence A. Welsch and Richard M. Cyert.* \$3.50

PROGRAMMING FOR OPTIMAL DECISIONS. *Edited by P.G. Moore and S.D. Hodges.* \$2.95

MARKETING RESEARCH. *Edited by Gordon Wills.* \$2.65

For a complete list of Penguin books in economics and business management, write:



PENGUIN BOOKS INC
7110 Ambassador Rd Baltimore, Md. 21207

Overseas readers, write: Penguin Education, Harmondsworth, Middlesex, England

For 1971

NEW ECONOMICS BOOKS FROM IRWIN

FORECASTING METHODS

By Roger K. Chisholm, Northwestern University, and Gilbert R. Whitaker, Jr., Washington University (St. Louis)

Stressing methodology, this basic technique-oriented text describes opinion or intentions surveys, economic indicators, and opportunistic forecasting methods, as well as quantitative forecasting methods including use of input-output and regression analysis.

About 350 pages / paperback

MONEY AND BANKING

By Douglas Fisher, University of Essex (England)

Taking a theoretical approach, this text places the money market in its proper place at the base of a pyramid of closely interrelated financial markets. The book presents a modern view of traditional money and banking material with emphasis on money as wealth.

About 500 pages

MANPOWER ECONOMICS: The American Labor Market in Action

By Lowell E. Gallaway, Ohio University

Blending conventional economic theory and empirical analysis, this text systematically investigates manpower and labor market problems. Topics covered include composition and variations in the aggregate labor force, labor mobility, unemployment-inflation trade-off, union impact, hidden unemployment, and the labor market's contribution to urban-rural economic problems.

About 350 pages / paperback

LABOR AND INDUSTRIAL SOCIETY, Third Edition

By Abraham L. Gitlow, New York University

This Third Edition is thoroughly reworked and updated and focuses more sharply on the impact of industrialization on wage determination, unemployment and trade unionism.

About 650 pages

MODERN PUBLIC FINANCE, Revised Edition

By Bernard P. Herber, University of Arizona

The Revised Edition of this basic text has been updated so as to incorporate recent developments in the field of public finance. Some of the principal changes include expanded treatment of welfare politics, principles of tax equity, the new federal unified budget concept, trade-off between full employment and price stability, and the public sector.

About 500 pages

MACROECONOMICS: Fluctuations, Growth, and Stability, Fifth Edition

By Maurice W. Lee, University of North Carolina

Stressing the imperative need to improve macroeconomic efficiency to meet the problems now challenging the economy, the Fifth Edition is designed to promote understanding and enable the reader to deal with the primary issues of the macroeconomy in the early 1970's.

About 700 pages

For examination copies, please write on departmental stationery indicating course title and text presently used for each book you request.

RICHARD D. IRWIN, INC.

Homewood, Illinois 60430

For 1971

NEW ECONOMICS BOOKS FROM IRWIN

WELFARE AND COMPETITION, Revised Edition

By Tibor Scitovsky, Stanford University

A new section entitled "Irrationality, Indivisibility, Inequality" has been added to this new edition. This part deals with public goods, pricing in state-owned enterprises, externalities, and the problems of the environment as well as those topics listed in the title of the section. About 500 pages

ECONOMICS OF FINANCIAL INSTITUTIONS AND MARKETS

By Paul F. Smith, Wharton School of Finance and Commerce, University of Pennsylvania

Taking a theoretical approach, this text integrates materials on the nature of financial decisions with micro- and macroeconomic theory. It views the financial system comprehensively, with emphasis on the behavior of economic units and the interaction between real and financial forces. About 250 pages

INTRODUCTION TO INTERNATIONAL ECONOMICS, Fifth Edition

By Delbert A. Snider, Miami University (Ohio)

This revised and updated edition features a complete re-writing of the section dealing with theory of trade. The Heckscher-Ohlin theorem is treated as a special case, and a more general approach is taken. About 500 pages

MONEY AND MARKETS

By Beryl W. Sprinkel, Senior Vice President, Harris Trust and Savings Bank, Chicago

Based on the author's earlier book, *MONEY AND STOCK PRICES*, this new volume discusses the monetarists and the new economists, money and inflation, money and the bond market, and monetary policy for the 1970's. About 300 pages

SELECTED ANTITRUST CASES: Landmark Decisions, Fourth Edition

By Irwin M. Stelzer, President, National Economic Research Associates

This new edition represents an updating of previous editions and consists of a compilation of landmark judicial decisions in the field of antitrust law. The cases selected promote understanding of the law as it now stands and knowledge of the historical development of the attitude of the courts on the various legal-economic issues. About 300 pages / paperbound

PUBLIC POLICIES TOWARD BUSINESS, Fourth Edition

By Clair Wilcox, late of Swarthmore College

Discussing the major types of public policy toward business activity, this new edition covers the control of monopoly by maintaining competition, by regulation and through public enterprise, the control of competition to protect consumers, investors and future generations, and to serve the interests of organized producer groups. About 850 pages

For examination copies, please write on departmental stationery indicating course title and text presently used for each book you request.

RICHARD D. IRWIN, INC.

Homewood, Illinois 60430



FINANCIAL MANAGEMENT TEXTS BY JAMES VAN HORNE

NEW EDITION—comprehensive, up-to-date introduction to the field of finance...

FINANCIAL MANAGEMENT AND POLICY, 2ND EDITION, 1971—James Van Horne, Stanford University

Significant portions have been changed in this revision, but its purpose remains: to develop an understanding of financial theory in an organized manner; to become familiar with the application of analytical techniques to a number of areas of financial decision making; and to expose the student to the institutional material necessary to give him a feel for the environment in which financial decisions are made.

Van Horne has reflected the changes that have occurred in financial theory and practice since the first edition, as well as sharpening and updating existing material so that it is better structured and more easily comprehended. There is an increased emphasis upon linking various financial decisions with valuation.

- Requires only background in elementary algebra and statistics; more difficult mathematics are treated in appendices.
- Concise coverage of institutional material.
- Suited in length for a semester-to-year course—608 pages. Spring 1971, 608 pp., cloth/\$11.95 (31530-9)

Representing a reformulation of the companion text, **FINANCIAL MANAGEMENT AND POLICY**, and concentrating on basic issues . . .

FUNDAMENTALS OF FINANCIAL MANAGEMENT

Van Horne provides a conceptual orientation intertwined with the institutional material needed to give the student a proper grounding in finance. He imparts a basic understanding of theory and practice as well as an awareness of changes that have occurred, and are presently occurring in the field.

Theory discussed is of an elementary nature. Emphasis is upon financial decision making and the concentration is on principles rather than upon analytical techniques.

- Undergraduate level: for those who liked **FINANCIAL MANAGEMENT AND POLICY 1** but found it too difficult.
- No advanced quantitative techniques—all the mathematical appendices removed.
- More complete institutional material for a more descriptive course.
- Designed for a quarter-to-semester course—512 pages. March 1971, 512 pp., cloth/\$10.95 (33924-2)

For more information write **Box 903**

PRENTICE-HALL

englewood cliffs, n.j. 07632

Orders will be processed faster if title and title codes appear on your order. Prices slightly higher in Canada.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

**Economic realities do not take place in a vacuum.
Economic studies show their diverse interrelationships.**

LINEAR ECONOMIC THEORY

Daniel C. Vandermeulen, Claremont Graduate School

This text offers the most thorough exploration of the relationship between linear programming and economic theory yet to appear. There are no mires of technical detail or computational devices to bog down readers primarily concerned with economic theory. The book is written cleanly and crisply for senior and graduate courses in linear economic theory and as a graduate level supplement in micro-theory for economics curricula or MBA programs.

May 1971 approx. 528 pp. \$10.95 (53689-6)

DEMAND, SUPPLY AND THE MARKET MECHANISM

Richard E. Hatwick, Western Illinois University

Joel W. Sailors, University of Houston

Bernard Brown, Sangamon State University, Springfield, Illinois

A new text, DEMAND, SUPPLY AND THE MARKET MECHANISM restricts its analytical apparatus to simple supply and demand diagrams, stressing the interrelatedness of markets. It gives students some unifying principles in economics for easier, more interesting and more successful study. The text analyzes the multiple and often conflicting economic objectives pursued by market participants. It stresses dynamic adjustment processes, and analyzes, along with market solutions, non-market solutions of economic decisions regarding public goods. The exposition of the text is designed to appeal to the novice as well as to the graduate student. By moving on two levels it explains theory satisfactorily for the undergraduate while offering an integrating, non-mathematical framework to the specialist.

April 1971 144 pp. \$3.95 (19799-6)

ECONOMIC DEVELOPMENT OF THE NORTH ATLANTIC COMMUNITY

Dudley Dillard, University of Maryland

ECONOMIC DEVELOPMENT OF THE NORTH ATLANTIC COMMUNITY traces the growth of the European and American economies from medieval origins to the present. The author views the evolution of these two economies as two closely related phases of a common development. Professor Dillard introduces basic ideas in economic analysis to give a broader insight into the organization and function of the North Atlantic economic system. The text also presents the classical theory of distribution, relating wages, profits and rent, to show the significance of free trade to England's industrial leadership during the nineteenth century. Professor Dillard employs the macroeconomic theory associated with the English economist John Maynard Keynes to explain the behavior of the North Atlantic economic system as a whole. Keynes' concepts of investment, saving, consumption and aggregate income are particularly relevant, the author notes, in view of the trend of European and American capitalism since the Great Depression.

1967 747 pp. \$12.50 (22330-5)

For further information about any of these books, write: Box 903

PRENTICE-HALL

Englewood Cliffs, New Jersey 07632

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

BACH'S ECONOMICS

**STUDENTS WANT TO PUT THE WORLD TOGETHER
IN A BETTER WAY.**

ECONOMICS: An Introduction to Analysis and Policy
Seventh Edition, 1971
GEORGE LELAND BACH
Stanford University

**NEW MATERIAL APPLIES ECONOMIC THEORY
TO THE BIG ISSUES OF THE DAY:**

Pollution and the Environment
The Population Explosion
Economic Growth and the "Quality of Life"
Poverty and Insecurity
Economics of the Draft
The Military/Industrial Complex
The Underdeveloped Economies
Black Economic Development and Black Capitalism
Marx, Radical Economics, and Profits
Economics of Minimum Wage Laws
The Urban Crisis

THOROUGHLY REVISED THROUGHOUT. Contains an especially thorough and up-to-date treatment of money and modern monetary analysis, reflecting new developments in the areas of money and monetary policy. Includes empirical evidence on both monetary and fiscal policy effectiveness, and provides a modern integration of monetarist and income-flows approaches to macro-analysis.

A COMPLETE COURSE IN ITSELF. Incorporation of new material now makes this text a complete course in itself. However, for teachers who want an expanded approach—the new Bach is again designed to fit well with a book of readings and outside material.

NEW LARGE QUESTION BANK. The best available with any elementary text, this question bank will be provided to adopters. Questions have been designed with great care and concern for educational principles of optimal test construction. Sufficient questions are available for both tests and daily class use.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

THE NEW EDITION OF BACH'S ECONOMICS CAN HELP THEM UNDERSTAND WHAT'S INVOLVED.

SUPPLEMENTS

ECONOMIC ANALYSIS AND POLICY: BACKGROUND READINGS FOR CURRENT ISSUES, 3rd Edition, 1971, Edited by Myron L. Joseph, Carnegie-Mellon University, Norton C. Seeber, Oakland University, and George Leland Bach

Readings in depth that focus directly on conflicting statements by economists, political leaders, and other observers on major current issues stressed by the text. Spring 1971.

TRANSPARENCIES, A set of 35 overhead transparencies, featuring multicolored overlays

INSTRUCTORS' MANUAL AND QUESTION BANK, G. L. Bach and William Moffat, Stanford University

WORKBOOK IN ECONOMICS, 5th Edition, 1971, Myron L. Joseph, Carnegie-Mellon University; and Norton C. Seeber, Oakland University

MACROECONOMICS: A PROGRAMMED BOOK, 2nd Edition, 1970, Richard Attiyeh, University of California, San Diego; Keith Lumsden and G. L. Bach, both at Stanford University

MICROECONOMICS: A PROGRAMMED BOOK, 2nd Edition, 1970, Keith Lumsden and G. L. Bach, both at Stanford University; Richard Attiyeh, University of California, San Diego

**For further information
write Box 903**

**PRENTICE-HALL
Englewood Cliffs, N.J. 07632**

7TH

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

The PRENTICE-HALL Series in **MATHEMATICAL ECONOMICS**

The Series is intended as a vehicle for making mathematical reasoning and quantitative methods available to the main corpus of the undergraduate and graduate economics curricula.

DONALD V. T. BEAR, University of California, La Jolla, Series Editor

LINEAR ALGEBRA FOR SOCIAL SCIENCES

Menahem E. Yaari, The Hebrew University, Jerusalem

Designed for the advanced social science student, this text helps the student gain a familiarity with the basic modes of thinking involved in mathematical reasoning. Emphasis is placed on the structures and concepts that provide a basis for many of the theoretical and statistical models used by the social scientist today.

January 1971 192 pp. \$9.95 (53683-9)

MATHEMATICAL OPTIMIZATION AND ECONOMIC THEORY

Michael D. Intriligator, University of California, Los Angeles

This self-contained text is distinctive in that it covers both programming and control techniques, emphasizing their important interrelationships and the logic of their development. The mathematical level is kept as elementary as possible—a basic knowledge of analysis and matrix algebra is sufficient. Designed as a text for mathematical optimization or economics, the format provides a strong supplementary background for courses in economic theory and operations research.

January 1971 544 pp. \$13.95 (56175-3)

BASIC ECONOMETRICS

Dennis Aigner, University of Wisconsin, Madison

BASIC ECONOMETRICS stresses the theory and application of regression methods in econometrics. The book features problems as an integral part of the format and focuses on applications of methods to economic models. A knowledge of probability and statistics is required.

March 1971 approx. 256 pp. \$10.95 (05901-4)

ELEMENTARY MATHEMATICAL MACROECONOMICS

David A. Bowers and **Robert N. Baird**, both of Case Western Reserve

A class-tested macroeconomics text, features a format that enables the instructor to cover most of the basic issues in contemporary macroeconomics in a minimum of time. The text begins at a low level of mathematical sophistication and never goes beyond multivariate calculus. The economic theory involved, however, becomes fairly demanding as the text progresses.

May 1971 approx. 320 pp. \$9.95 (25832-7)

MATHEMATICAL TECHNIQUES FOR ECONOMIC ANALYSIS

Arthur Benavie, University of North Carolina, Chapel Hill

This new book develops the foundations of static and economic analysis. The mathematical techniques that are indispensable to the economic theorist are spelled out with rigor and are given an intuitive basis with lucid explanations. The book is accessible to the student who has had mathematical economics or the equivalent in mathematics courses along with macroeconomic or microeconomic theory.

Forthcoming

A MATHEMATICAL BACKGROUND FOR ECONOMISTS AND SOCIAL SCIENTISTS

Ronald C. Read, University of Waterloo, Ontario

Professor Read provides a thorough background in those branches of mathematics that are increasingly being used in economics and social sciences. The author also explores more specialized topics of interest to the economist—including linear programming, game theory, graph theory, and computer programming. Designed as a course text, the style and presentation make it an ideal book for the student involved in independent study.

July 1971 approx. 848 pp. \$15.95 (56098-7)

For further information about any of these books, write: Box 903

PRENTICE-HALL

Englewood Cliffs, New Jersey 07632

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

From the Prentice-Hall
Foundations of Modern Economics Series

INTERNATIONAL ECONOMICS, 3rd Edition, 1971

by Peter B. Kenen and Raymond Lubitz, both of Columbia University

The 3rd Edition of INTERNATIONAL ECONOMICS provides a rigorous and concise introduction to international economics. It updates and expands issues on trade problems, recent financial development, current U.S. balance-of-payments problems, and the creation of Special Drawing Rights. It also analyzes trade policy from a historical as well as a contemporary point of view and thoroughly examines international trade theory—comparative advantage, the gains from trade and the effects of tariffs. Chapter 6 discusses trade policies toward less-developed countries, especially tariff preferences and the theory of effective protection. February 1971, 144 pp., illus., paper \$2.75 (47261-3)

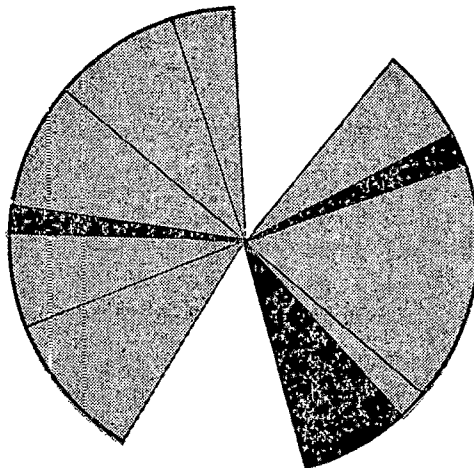
NATIONAL INCOME ANALYSIS, 3rd Edition, 1971

by Charles L. Schultze, Brookings Institute

Featuring a complete introduction to economics as a problem of choice and priority setting, the 3rd Edition analyzes factors that determine the overall levels of economic activity, national income, employment, and prices. This edition opens with an introduction to the role of national income analysis in economics. It also discusses the measurement of national income and Gross National Product, and then explains in narrative, graphical and algebraic forms the fundamental economic relationships that determine their magnitude. A brief discussion follows of how monetary and fiscal policy affect these relationships and thereby influence aggregate output. May 1971, approx. 176 pp., paper \$2.95 (60942-0)

for more
information,
write box 903

Prentice-Hall,
Englewood Cliffs,
N. J. 07632



From Business Publications, Inc.

NOTEWORTHY TEXTS IN ECONOMICS

MANAGERIAL ECONOMICS: Analysis and Cases, Revised Edition

By W. Warren Haynes, State University of New York

This recently published basic text for managerial and business economics courses stresses the application of certain fundamental economic concepts to managerial macro-problems. An Instructor's Manual is available.

740 pages / \$11.50 / 1969

MONEY, BANKING, AND THE NATION'S INCOME

By David T. Lapkin, University of North Carolina

Treating the economics of money and banking in a concise, nonencyclopediaic manner, this volume covers the creation and destruction of money, the structure and function of commercial banking and the Federal Reserve System, and the effects of economic factors upon bodies implementing monetary policy. A Teacher's Manual is available.

432 pages / \$9.50 / 1969

PUBLIC FINANCE: An Introduction to the Study of the Public Economy, Revised Edition

By Ansel M. Sharp, Oklahoma State University, and Bernard F. Sliger, Louisiana State University

Divided into four major subject areas, this edition places emphasis on fiscal policy and theory and on current public finance problems such as metropolitan needs and the finance of public education and highways.

430 pages / \$10.50 / 1970

LABOR ECONOMICS, Third Edition

By Chester A. Morgan, University of Iowa

This text gives a chronological presentation of the institutions and problems of the labor market, as well as the latest legislative and court developments related to the labor market as they have emerged in the industrial society.

675 pages / \$11.50 / 1970

THE NATURE OF PRICE THEORY, Revised Edition

By H. H. Liebhafsky, University of Texas

This Revised Edition gives an integrated and comprehensive coverage of basic concepts and relationships in price theory, explaining concepts verbally, algebraically, graphically, and by use of numerical equivalents inserted into simple algebraic equations.

662 pages / \$11.50 / 1968

For examination copies, please write on departmental stationery indicating course title and text presently used for each book you request.

Business Publications, Inc.
3511 Manor Road
Austin, Texas 78723



Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

2 New Texts on

WELFARE POLICY AND INDUSTRIALIZATION IN EUROPE, AMERICA AND RUSSIA

By GASTON V. RIMLINGER,
Rice University

Rimlinger, an economist with a strong background in sociology, traces the evolution of the individual's right to income security and health in industrialized countries.

He compares England, France, and the United States where self-help tradition has hampered broad state-help programs to Germany where a strong patriarchal protection is predominant, and to Russia where collectivism has replaced patriarchalism.

And he shows how each country arrived at solutions suited to its own particular circumstances.

In this sensitive interpretation of these ideologies, he raises such questions as: to what extent should the individual be held responsible for his own welfare or to what extent should he rely on society? How economic protection affects the individual's freedom, initiative, and foresight? And what are the political and economic consequences of various kinds and levels of protection?

The answers to these questions depend on the interplay of economic, social, and political factors. And on how those in power view protection and intervention of government.

1971 Approx. 384 pages In press

ECONOMIC GROWTH AND DEVELOPMENT:

A Mathematical Introduction

By PHILIP A. NEHER, University of British Columbia, Vancouver

An introduction to the mathematical theories of economic growth and development for the mathematically weak economist.

The book presents the core of modern growth and development literature on two levels: (1) Assuming the student doesn't know calculus. And (2) assuming that he does.

Among the highlights of this text are the analysis of the growth of an international economy (the first in a book on this level), and the exploration of the neoclassical production function, complete with comprehensive tables that summarize the more important relations.

1971 352 pages \$9.95

For more information, contact your local Wiley representative or write Wayne Anderson, Promotion Manager.



JOHN WILEY & SONS, Inc.

605 Third Avenue, New York, N.Y. 10016

In Canada:

22 Worcester Road, Rexdale, Ontario

HRW MAKES ONE THING PERFECTLY CLEAR— ECONOMICS

PRINCIPLES OF ECONOMICS

By Donald A. Nichols, University of Wisconsin, and Clark W. Reynolds, Stanford University

This new and comprehensive text forms the basis for analysis in basic economics courses. The authors discuss the contributions of the historical and modern leaders in the field and economic systems of various countries. The student is guided from the elements of the theories of production and demand to general equilibrium, and then to income determination, trade, growth, and development. Considerable material new to principles texts, but essential to the student's understanding of modern economics, is included.

March 1971 / 704 pages / \$9.95 (tent.)
Instructor's Manual

Study Guide for Principles of Economics by Charles Metcalf, University of Wisconsin
May 1971 / 288 pages / \$3.95 paper (tent.)

READINGS FOR PRINCIPLES OF ECONOMICS

Edited by Donald A. Nichols and Charles Metcalf, both of the University of Wisconsin

May 1971 / 640 pages / \$5.95 paper (tent.)

THE ACADEMIC SCRIBBLERS: American Economists in Collision

By William Breit, University of Virginia, and Roger Ransom, University of California, Riverside

February 1971 / 256 pages / \$8.50 cloth
\$3.95 paper

TANSTAAFL: The Economic Strategy for Environmental Crisis

By Edwin G. Dolan, Dartmouth College

April 1971 / 160 pages / \$2.50 paper

A RADICAL'S GUIDE TO ECONOMIC REALITY

By Angus Black

1970 / 87 pages / \$1.95 paper

ESSENTIAL PRICE THEORY

By Albert Levenson and Babette Solon, both of Queens College, The City College of New York

January 1971 / 352 pages / \$8.95

STATISTICS AND ECONOMETRICS

By Charles R. Frank, Jr., Princeton University

April 1971 / 320 pages
\$10.00 (tent.) *Economics Series*

INTRODUCTION TO MICROECONOMICS

By Richard H. Leftwich, Oklahoma State University

1970 / 384 pages / \$4.15 paper

READINGS IN MACROECONOMICS

Second Edition

Edited by M. G. Mueller, University of Glasgow

May 1971 / 472 pages / \$6.95 (tent.)

**READINGS IN
MICROECONOMICS**

Second Edition

Edited by William Breit, University of Virginia, and Harold Hochman, The Urban Institute, Washington, D.C.

June 1971 / 544 pages / \$6.95 paper (tent.)

READINGS IN LABOR MARKET ANALYSIS

Edited by John F. Burton, Jr., Lee K. Benham, William M. Vaughn, III, and Robert J. Flanagan, all of the University of Chicago

May 1971 / 500 pages / \$5.95 paper (tent.)

ESSENTIALS OF MANAGERIAL FINANCE Second Edition

By J. Fred Weston, University of California, Los Angeles, and Eugene F. Brigham, University of Wisconsin

March 1971 / 544 pages / \$10.50
Instructor's Manual

Study Guide for Essentials of Managerial Finance, Second Edition; Managerial Finance, Third Edition

April 1971 / 208 pages / \$4.15 paper

AND NOW—60 TRANSPARENCIES FREE TO ADOPTERS

READINGS IN MANAGERIAL FINANCE

Edited by Eugene F. Brigham, University of Wisconsin

April 1971 / 320 pages / \$4.95 (tent.)

STATISTICS: Probability, Inference and Decision (Volumes I and II now combined in one volume)

By William Hays, University of Michigan, and Robert Winkler, Indiana University

January 1971 / 960 pages / \$13.95
Solutions Manual

Still available in individual volumes, I and II
Quantitative Methods for Decision-Making Series

ELEMENTS OF PROBABILITY AND STATISTICS

By Steven A. Lippman, University of California, Los Angeles

March 1971 / 320 pages / \$10.95 (tent.)

Instructor's Manual

Quantitative Methods for Decision-Making Series

AN INTRODUCTION TO PROBABILITY, DECISION AND INFERENCE

By Irving H. LaValle, Tulane University

1970 / 784 pages / \$14.95

Quantitative Methods for Decision-Making Series

INTRODUCTION TO LINEAR PROGRAMMING

By Chaiho Kim, University of Santa Clara

February 1971 / 384 pages / \$13.95 (tent.)



**Holt, Rinehart,
and Winston, Inc.
383 Madison Ave.
New York, N.Y. 10017**

KNOPF—The Alternative

The Wiley INTRODUCTION TO ECONOMIC SERIES, edited by Kenyon A. Knopf, answers the need for a flexible alternative to large omnibus textbooks. This series is that alternative—it provides a solid core of economic principles while permitting instructors the flexibility of introducing a select set of problems and applied ideas.

The Latest Two from the Knopf Series—

1. THE ECONOMICS OF POVERTY, Second Edition

By ALAN B. BATCHELDER, Kenyon College

The unique feature of this book is that it leaves decisions to the reader. While it outlines the traits, causes and possible remedies of persistent poverty in America.

Earlier discussions of the extent and etiology of American poverty are expanded. Charities, social security and other transfers of money from the non-poor to the poor are examined. And positive alternatives out of the impasse are provided.

1971 (Approx.) 272 pages cloth: \$6.95
paper: \$2.95

2. CASE STUDIES IN AMERICAN INDUSTRY, Second Edition

By LEONARD W. WEISS,
University of Wisconsin

The core of this valuable text-supplement is 5 case histories, authored by Weiss to update his widely acclaimed first edition.

Studies on agriculture, electric power, steel, retailing and steel labor.

Each contributes to an understanding of micro-economic theory and policy; shows the structure and performance of a vital industry or factor market; and illustrates applications of economic theory.

Plus. Public policy with regard to a slew of issues—including public ownership of utilities, the collusion, merger and "monopolizing" rules of antitrust in steel, the Robinson Patman Act, cartels, and trade unionism—is presented and analyzed.

Best, the author achieves generality by surveying competition in product and factor markets at the outset; and literature on industry performance at the conclusion.

1971 (Approx.) 384 pages cloth: \$6.95 (tent.)
paper: \$3.95 (tent.)

INTRODUCTION TO ECONOMICS SERIES

Edited by KENYON A. KNOPF,
Jentzen Professor of Economics, Grinnell College

Other volumes in the Series

THE ECONOMICS OF THE PUBLIC SECTOR

By Robert H. Haveman, Grinnell College
1970 225 pages cloth: \$6.95 paper: \$2.95

NATIONAL INCOME AND EMPLOYMENT ANALYSIS, Second Edition

By Arnold Colclery, Amherst College
1970 162 pages cloth: \$5.95 paper: \$2.95

THE MARKET SYSTEM, Second Edition

By Robert H. Haveman and Kenyon A. Knopf,
both of Grinnell College
1970 269 pages cloth: \$6.95 paper: \$2.95

ECONOMIC DEVELOPMENT AND GROWTH

By Robert E. Baldwin, University of Wisconsin
1966 133 pages cloth \$6.50 paper: \$3.50

INTERNATIONAL ECONOMIC PROBLEMS, Second Edition

By James C. Ingram, University of North Carolina
1970 182 pages cloth: \$6.95 paper: \$2.95

TOWARD ECONOMIC STABILITY

By Maurice W. Lee, University of North Carolina
1966 177 pages cloth: \$6.50 paper: \$2.50

WELFARE AND PLANNING: AN ANALYSIS OF CAPITALISM VS. SOCIALISM

By Heinz Kohler, Amherst College
1966 175 pages cloth: \$6.50 paper: \$3.50

For further information, contact your local Wiley representative, or write Wayne Anderson, Promotion Manager.



JOHN WILEY & SONS, Inc.
605 Third Avenue, New York, N.Y. 10016

In Canada:
22 Worcester Road, Rexdale, Ontario

COLUMBIA UNIVERSITY PRESS

The Widening Gap:

DEVELOPMENT IN THE 1970's. Barbara Ward, Albert S. Schweitzer, with the assistance of Lee d'Anjou and J. D. Rumlalls. A selection of the most important of the papers commissioned by the Columbia University Conference on International Development in response to the report of the Pearson Commission. Their virtually unanimous conclusion is that the 1970's will be a period of deepening social and economic strain in the relations between developed and developing countries and also within the developing nations themselves.

\$10.00 cloth
\$3.95 paper

The Inter-War Economy: Britain, 1919-1939

Derek H. Aldcroft. Avoiding the standard factual and descriptive approach, this volume focuses on the major themes and issues of debate during the period covered. The performance of the economy is assessed in historical perspective and against the background of international economic influences.

\$12.50

Government Enterprise

Wolfgang G. Friedmann and J. F. Garner. Incorporates the results of a colloquium held to make a comparative study of the development of government holding companies and governmental shareholding in private companies. Great Britain, France, Italy, Germany, Sweden, the United States, Canada, Australia, Israel, and East Africa are among the countries covered.

\$10.00

Joint International Business Ventures In Developing Countries: Case Studies and Analysis of Recent Trends

Wolfgang G. Friedmann and John Pierre Beguin. Analyzes in depth some 20 cases of joint ventures in which developed and developing countries are associated. The case studies have been chosen so as to cover the widest possible variety of both bi-partite and multipartite joint ventures, and they include partnerships between private enterprises as well as government participations and mixed private-public associations.

\$15.00

The Tax System of Liberia:

REPORT OF THE TAX MISSION. Carl S. Shoup, Douglas Dosser, Rudolph G. Penner, and William S. Vickrey. The full text of the Report to the Liberian Government on that country's tax system, prepared during the summer of 1969. It presents the latest data on the size and structure of the national income of Liberia as background for the description and analysis of each of the Liberian taxes, followed in each case by specific recommendations for change.

\$8.50

The Regulation of the Consumer Finance Industry

Michael Kawaja. This study seeks to answer two questions: How did the rate and loan size regulation in the consumer finance industry evolve? What are their effects on the character of the lending service provided, the volume of lending facilities that are supplied under regulation, and the earnings rates of lenders? New York State is used as the base for this study and a comparison is made of data available from other states.

Studies in Consumer Credit, No. 3, Graduate School of Business, Columbia University.

\$4.00

Address for orders: 136 So. Broadway, Irvington, N. Y. 10533
In Canada: McGill-Queen's University Press, 3458 Redpath St., Montreal 109

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

FROM McGRAW-HILL

ECONOMICS AND SOCIAL CHOICE

Jack W. Nickson, Jr., Old Dominion University. 288 pages, \$5.95 soft. Study Guide: \$2.95 soft

The purpose of this text is to explain *how* and *why* the economy performs as it does. More than other books in the field it stresses current social and economic problems and suggests specific remedies for these problems.

NATIONAL INCOME AND THE PRICE LEVEL: A Study in Macroeconomic Theory, Second Edition

Martin J. Bailey, University of Rochester. 288 pages, \$10.95

Professor Bailey provides a clear and concise presentation of the theory of national income determination and of some of the problems of theory and policy that flow directly from it. Concentrating on basics, he treats theory first and then develops ramifications and applications. There are new chapters on inflation and growth and a reorganized treatment of consumption, investment, and money.

PROBLEMS IN MICROECONOMICS

Richard A. Bilas, California State College, Bakersfield, and **Richard S. Wallace**, Winthrop College. 272 pages, \$5.95 soft

This workbook is designed as a teaching aid for professors of microeconomic theory. By solving the problems contained in it the student gains a better understanding of the fundamentals of price theory as well as an appreciation of some of the subtleties of the subject.

PERSPECTIVES IN ECONOMICS

Edited by **Alan A. Brown**, Indiana University, **Egon Neuberger**, State University of New York at Stony Brook, and **Malcolm Palmatier**, The RAND Corporation. 272 pages, \$4.50 soft

Here is a collection of class-tested essays contributed by eminent economists. It may be used as core reading or as a supplement to college courses in the principles of economics. The volume presents a clear perspective of the various fields of economics and their relation to the discipline as a whole.

THE ECONOMICS OF CORPORATION FINANCE

Sergei P. Dobrovolsky, Rensselaer Polytechnic Institute. 488 pages (tent.), \$13.95 (tent.). Available Spring 1971

The purpose of Professor Dobrovolsky's book is to provide a theoretical framework for analyzing business financial policies and problems. The first part is concerned with microfinance, namely, the financial operations of an individual firm, particularly, a corporate enterprise, while the second part deals with macro-finance, that is, the capital market, flow of capital funds, and determination of the aggregate volume of business investment.

A SURVEY OF GENERAL EQUILIBRIUM SYSTEMS

Bent Hansen, University of California, Berkeley. *Economics Handbook Series*. 256 pages, \$9.95

The author rigorously builds the structure of general equilibrium systems and carries the analysis up to modern equilibrium models. The text lucidly and impartially presents essential issues of monetary theory, unemployment and inflation, international trade, and growth.

McGraw-Hill Book Company • 330 W. 42nd Street • New York, N.Y. 10036

MICROECONOMIC THEORY: A Mathematical Approach, Second Edition

James M. Henderson, University of Minnesota, and **Richard E. Quandt**, Princeton University. *Economics Handbook Series*. 448 pages, \$12.95

Offering a comprehensive treatment of microeconomic theory that uses a classical mathematical approach, this edition covers the theory of the firm, the theory of single market equilibrium, general equilibrium, welfare economics, capital theory, monopolies and oligopolies, and linear programming and related approaches.

THE ECONOMICS OF STATE AND LOCAL GOVERNMENT

Werner Z. Hirsch, University of California, Los Angeles. *Economics Handbook Series*. 333 pages, \$12.95

Written by a highly regarded economist, this book represents an important contribution to the study of public economics. It applies contemporary economic approaches to an analysis of the activities of the state and local sector.

ECONOMETRIC METHODS, Second Edition

J. Johnston, University of Manchester, England. 500 pages (tent.), \$11.00 (tent.). Available Spring 1971

Professor Johnston's text still retains the same operational and practical approach. Important features include a survey of the latest Monte Carlo results on simultaneous equation estimators, a special chapter on autocorrelation, and one on dummy variables and their applications in seasonal adjustment and the analysis of covariance.

MONETARY ECONOMICS: Readings on Current Issues

Edited by **William E. Gibson**, University of California, Los Angeles, and **George G. Kaufman**, University of Oregon. 450 pages (tent.), \$7.95 (tent.). Available Spring 1971

The collection of articles in this book joins together the issues under debate and subjects them to rigorous empirical analysis through the use of the scientific method and new statistical techniques.

ECONOMICS AS A SCIENCE

Kenneth E. Boulding, University of Colorado. 160 pages, \$6.95 hard; \$3.95 soft

This work provides introductory economics students with a perspective of the scope and setting of economics and the place it occupies in the intellectual community.

BASIC ECONOMIC PRINCIPLES, Second Edition

Robert A. Lynn, Kansas State University. 384 pages, \$8.50

The purpose of this one-term text is to: (1) present a survey of economic analysis, issues, and institutions; (2) relate this material to the important economic issues that are currently confronting the United States; and (3) familiarize the student with the basic tools that economists use to analyze economic situations. A Study Guide, Instructor's Manual, and Test File are also available.

McGraw-Hill Book Company • 330 W. 42nd Street • New York, N.Y. 10036

PERSPECTIVES ON WAGE DETERMINATION: A Book of Readings

Campbell R. McConnell, University of Nebraska. 352 pages, \$7.95 hard; \$5.95 soft

Here is a collection of materials surveying the main trends of thought on wage determination and depicting the current status of wage theory.

ARBITRATION AND COLLECTIVE BARGAINING: Conflict Resolution in Labor Relations

Paul Prasow and **Edward Peters**, both of the University of California, Los Angeles. *McGraw-Hill Series in Management*. 480 pages, \$12.95

The thesis of this book is that the dynamics of collective bargaining can best be understood by studying arbitral issues and concepts. Analytical concepts, rather than techniques, are stressed and a variety of actual case studies are included.

INDUSTRIALIZATION AND ECONOMIC HISTORY: Theses and Conjectures

Jonathan R. I. Hughes, Northwestern University. 300 pages, \$6.95

This book's general purpose is to examine the economic, social, and political effects that industrialization has had in certain countries since the 18th century.

ECONOMICS, Eighth Edition

Paul A. Samuelson, Massachusetts Institute of Technology. 868 pages, \$10.50

Study Guide (by Robinson): \$3.95. *Economics: An Introductory Program* (by Foster): \$4.95. *Readings in Economics*, 6/e: \$7.95 hard, \$5.95 soft. Instructor's Manual and Test Bank: Free to adopters. Visual Aids.

CONTEMPORARY ECONOMICS: Principles and Issues

Leonard S. Silk, The Brookings Institution. 500 pages, \$8.95

Analytical and problem-oriented in its approach, this text presents many real-world examples that are within the experience of the average student. *Readings in Contemporary Economics* by Leonard Silk is also available, as well as a Study Guide by Arthur L. Welsh, an Instructor's Manual, and a Test File.

THE LOGIC OF THE PRICE SYSTEM

Paul B. Trescott, Southern Methodist University. 450 pages, \$9.95

Utilizing a minimum of calculus, this book stresses the application of microeconomics to levels beyond the individual firm or consumer. The author elaborates on such economy-wide matters as factor pricing and income distribution, resource allocation, international trade, and economic growth.

McGraw-Hill Book Company • 330 W. 42nd Street • New York, N.Y. 10036

INTRODUCTORY ECONOMIC THEORY: An Audio-Visual-Tutorial Approach

David A. Martin, State University of New York at Geneseo. Complete set \$350 (tent.). Worktext \$3.95 (tent.). Instructor's Manual \$1.50.

1—Supply, Demand, and Equilibrium Price. 2—Macroeconomic Equilibrium. 3—The Investment Multiplier. 4—The Equalities of Saving and Investment. 5—The Creation of Money. Each unit \$75 (tent.)

INTRODUCTION TO CONTEMPORARY MICROECONOMICS

Vivian Walsh, University of Washington. 352 pages, \$8.95 hard; \$6.95 soft

After presenting a general theory of choice from microtheory, the author then develops models of consumption, production, and exchange. Elementary notions of first order logic, of naive set theory, and of certain topological properties are developed *within* the exposition.

PRINCIPLES OF FINANCIAL MANAGEMENT

Ward S. Curran, Trinity College, Connecticut. McGraw-Hill Series in Finance. 576 pages, \$10.50

This book utilizes the financial manager's viewpoint to examine the principles of microfinancial theory, including the management of working capital, fixed assets, and the composition of financing.

PORTFOLIO THEORY AND CAPITAL MARKETS

William F. Sharpe, University of California, Irvine. McGraw-Hill Series in Finance. 350 pages, \$12.95

This book is one of the first to offer a unified presentation of much of the material that has revolutionized the fields of investments, finance, and the economics of uncertainty. *Much of the material is the author's own work.*

ECONOMICS OF INTERNATIONAL TRADE

Richard I. Leighton, State University of New York at Binghamton. 240 pages, \$8.95

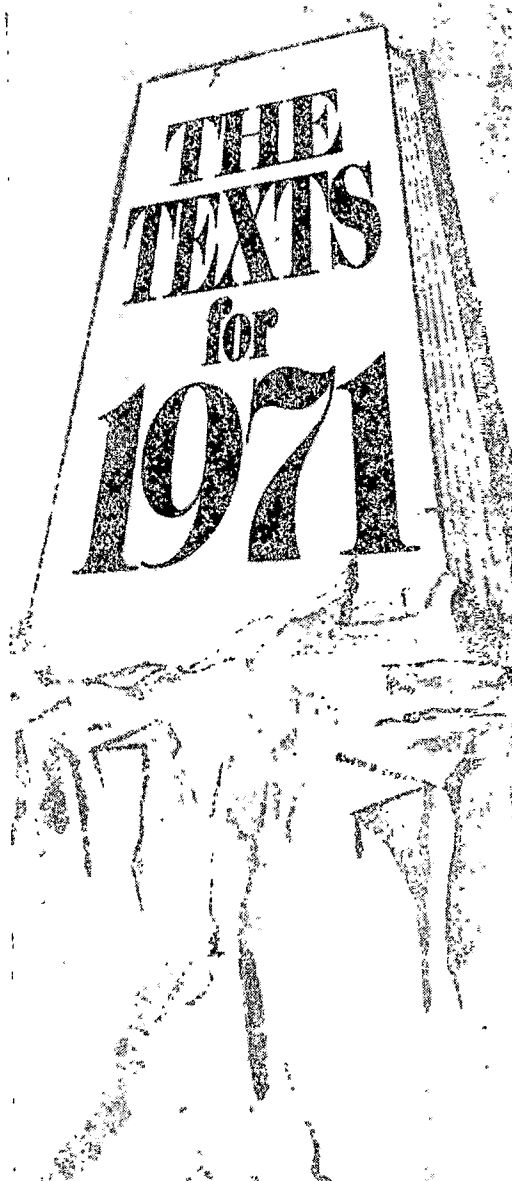
This short, theoretical book stresses both pure theory and monetary economics. The author examines the economic variables that determine the level of international commerce and their interaction with the domestic economy.

STATISTICAL ANALYSIS FOR MANAGERIAL DECISIONS

John C. G. Boot, State University of New York at Buffalo, and **Edwin B. Cox**, Boston University. 544 pages, \$11.95

This text has a dual purpose: (1) to present the essential elements and philosophy of business and economic statistics; and (2) to convey to the student an appreciation of the usefulness and decision-making power of statistics.

McGraw-Hill Book Company • 330 W. 42nd Street • New York, N.Y. 10036



THE TEXTS for 1971

PRINCIPLES OF ECONOMETRICS

By Henri Theil, The University of Chicago

Describes the techniques of handling statistical data in economics, with applications at every step. Designed both for students who want to be practitioners in data analysis and for whom this is the only econometrics course, and also for students who want to specialize in econometrics.

1971 736 pages \$14.95

COMPUTER SIMULATION EXPERIMENTS WITH MODELS OF ECONOMIC SYSTEMS

By Thomas H. Naylor, Social System Simulation Program, Duke University

This book is about the methodology for designing and implementing computer simulation experiments. Includes many example models taken from business and economics as well as descriptions of several complete simulation experiments. This is the only text on simulation that treats such topics as validation, stopping rules, sequential sampling, spectral analysis, multiple ranking procedures, and simulation versus analytical solutions.

1971 Approx. 512 pages \$12.95 tent.

REGRESSION AND ECONOMETRIC METHODS

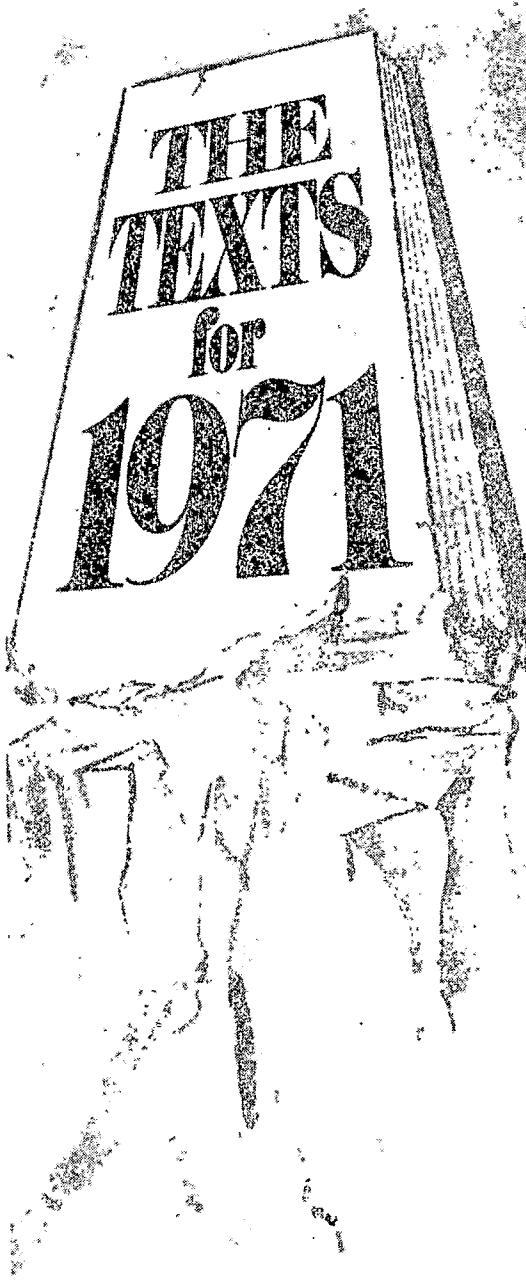
By David S. Huang, Southern Methodist University

Concerned with the principles and methods of least squares as they are employed in simple and multiple regression, tests of linear hypotheses, special econometric problems, multivariate regression, and simultaneous equation systems. The text gives an especially thorough treatment to linear restrictions, multicollinearity, and autocorrelation.

1970 274 pages \$12.95



JOHN WILEY & SONS, Inc.
605 Third Avenue, New York, N. Y. 10016
In Canada: 22 Worcester Road,
Rexdale, Ontario



THE TEXTS for 1971

THE NEW ECONOMIC HISTORY:

Recent Papers on Methodology

Edited by Ralph L. Andreano,
University of Wisconsin

A collection of some of the best papers written on the new quantitative techniques and methodologies of "The New Economic History." Articles by both proponents and skeptics are included and are arranged starting with pure methodological issues, followed by criticisms of method in substantive research, and finally a critique of the qualitative tools.

A volume in the Wiley Series in American Economic History

1970 178 pages Cloth: \$6.95 Paper: \$3.95

THE COURSE OF AMERICAN ECONOMIC GROWTH AND DEVELOPMENT

By Louis M. Hacker, Columbia University

An introduction to the forces and factors responsible for the economic growth and the resulting social and institutional development of the United States, historically viewed. The methodology employed is that of both the disciplines of economics and history. The book is designed, therefore, for either economics or American history students and for students of business administration.

A volume in the Wiley Series in American Economic History

1970 382 pages Cloth: \$3.95 Paper: \$4.95

LABOR ECONOMICS:

The Changing Labor Scene

By G. Glyn Williams, University of South Carolina

This book is an analytic survey of the American labor scene. Descriptive material, both institutional and statistical, is drawn from a wide variety of sources and is integrated into the book's analytical framework. The book describes, raises questions, discusses current research, and challenges the student to relate the facts to his own questions.

1970 489 pages \$9.95

MONEY AND BANKING:

An Introduction to Analysis and Policy

Second Edition

By John G. Ranlett, Sacramento State College

Provides a sound treatment of theory as a means of understanding and implementing economic policy with an expanded discussion of money supply theory. The student is introduced to the new liquidity theory of portfolio management and to money and bank credit supply functions.

1969 543 pages \$10.50

WILEY

JOHN WILEY & SONS, Inc.

605 Third Avenue

New York, N.Y. 10016

In Canada: 22 Worcester Road, Rexdale, Ontario

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

New Books

Operations Research



**Springer-Verlag
New York Inc.**

175 Fifth Avenue, New York,
New York 10010

Heertje, A., Grundbegriffe
der Volkswirtschaftslehre.
45 fig. 318 pp. 1970.
(Heidelberger Taschenbücher,
Vol. 78) DM 10,80; US \$ 3.00

Heidelberger Arbeitsbücher

Vol. 1: Schmid, B. A., Arbeits-
buch zu „Stobbe, Volkswirt-
schaftliches Rechnungswesen“.
137 pp. 1970.
DM 8,—; US \$ 2.20

Vol. 2: Zöller, W., Arbeitsbuch
zu Handelsbilanzen. 165 pp.
1970. DM 10,—; US \$ 2.80

Lecture Notes in Operations Research and Mathematical Systems

Vol. 30: Noltemeier, H.,
Sensitivitätsanalyse bei
diskreten linearen Optimie-
rungsproblemen. 108 pp. 1970.
DM 10,—; US \$ 2.80

Vol. 31: Kühlmeyer, M., Die
nichtzentrale t-Verteilung.
Grundlagen und Anwendungen
mit Beispielen. 108 pp. 1970.
DM 10,—; US \$ 2.80

Vol. 32: Bartholomes, F.,
Homomorphismen und Re-
duktionen linearer Sprachen.
155 pp. 1970.
DM 14,—; US \$ 3.90

Vol. 33: Hinderer, K., Founda-
tions of Non-stationary
Dynamic Programming with
Discrete Time Parameter.
166 pp. 1970. DM 16,—; US \$ 4.40

Vol. 34: Störmer, H., Semi-
Markoff-Prozesse mit endlich
vielen Zuständen. Theorie
und Anwendungen. 134 pp.
1970. DM 12,—; US \$ 3.30

Vol. 35: Ferschl, Markovketten.
174 pp. 1970. DM 14,—; US \$ 3.90

Vol. 36: Magill, M. J. P. On a
General Economic Theory of
Motion. 101 pp. 1970.
DM 10,—; US \$ 2.80

Vol. 37: Müller-Merbach, H.
On Round-Off Errors in
Linear Programming. 53 pp.
1970. DM 10,—; US \$ 2.80

Vol. 38: Statistische Metho-
den I. Grundlagen und Ver-
suchsplanung. Edited by
E. Walter. 346 pp. 1970.
DM 22,—; US \$ 6.10

Vol. 39: Statistische Metho-
den II. Mehrvariable Metho-
den und Datenverarbeitung.
Edited by E. Walter. 161 pp.
1970. DM 14,—; US \$ 3.90

Ökonometrie und Unternehmensforschung / Econometrics and Operations Research

Vol. XIV: Praktische Studien
zur Unternehmensforschung.
Von E. Nievergelt, O. Müller,
F. E. Schlaepfer, W. H. Landis.
252 pp. 1970.
DM 58,—; US \$ 16.00

Vol. XV: Müller-Merbach, H.,
Optimale Reihenfolgen.
45 fig. 234 pp.
DM 60,—; US \$ 16.50

Vol. XVI: Selten, R., Preis-
politik der Mehrprodukten-
unternehmung in der
statistischen Theorie. 20 fig.
203 pp. 1970.
DM 64,—; US \$ 17.60

Vol. XVII: Hyvärinen, L. P.,
Information Theory for
Systems Engineers. 44 fig.
Approx. 207 pp. 1970.
DM 44,—; US \$ 12.10

**Scientific Papers
of Tjalling C. Koopmans**
1 portr. 612 pp. 1970.
DM 54,—; US \$ 15.00

*Standing orders are invited for
our serial publications.*

BASIC'S ECONOMICS BOOKS ARE MORE THAN TIMELY



ECONOMIC HERESIES: Some Old-Fashioned Questions in Economic Theory

Joan Robinson

A great economist here attacks the new (Keynesian) orthodoxy and argues convincingly that it is nothing but the old neo-classical orthodoxy supported by Keynesian policies designed to keep investment at full employment levels. She then provides a new look at the theory of value, distribution, and accumulation—and proposes a fundamental reorientation of economic theory designed to clear away orthodox obfuscation. April/\$6.95

ECONOMIC ANTHROPOLOGY AND DEVELOPMENT

George Dalton

These fourteen seminal essays by one of this country's leading economic anthropologists provide a fascinating introduction to the entire subject—its scope, content, and methodologies, and how these relate to the problem of social change, development, and modernization. April/\$12.50

ON RELIEF: The Economics of Poverty and Public Welfare

Bruno Stein

The first attempt by an economist to present a comprehensive survey and critical analysis of the historical evolution, politics, and economic logic—or lack thereof—of public assistance to the poor. Prof. Stein finds such programs ill-adapted to cope with the needs of the poor, and proposes a new framework for the evaluation of public welfare and other assistance programs that realistically integrates income transfer proposals with the larger problem of poverty. June/\$7.50

MONEY: A Guide for the 1970's

Lawrence S. Ritter & William L. Silber

Now recognized as the standard text in the field, this book provides a unique guide to monetary theory and policy. "We have at last found the book which, in good conscience, instructors in money and banking can recommend to their students."—*Journal of Finance*. "Excellent . . . a joy to read."—Lester V. Chandler

Published/\$5.95

Paperback classroom edition/\$2.95

ECONOMICS AND URBAN PROBLEMS

Dick Netzer

One of the nation's top economists shows us how the tools of his trade can help us cope more rationally with problems as diverse as urban poverty and traffic congestion. In clear, non-technical language, Dr. Netzer discusses the kinds of questions economists ask and how these questions enable us to formulate problems in a more meaningful and operational fashion. Published/\$5.95

EXPLORATIONS IN MANAGERIAL ECONOMICS

Bela Gold

In this important volume, one of the pioneers of the "new" managerial economics presents workable "intermediate" analytical frameworks based on empirical industry data concerning productivity, costs, control ratios, and other factors. An essential supplement to existing textbooks as well as an indispensable guide in the field. May/\$8.95

BLACK BUSINESS ENTERPRISE: Historical and Contemporary Perspectives

Ronald W. Bailey, Editor

The first truly comprehensive sourcebook on the history, current status, and future prospects of black economic development. "Ronald Bailey has rendered a great service to everyone concerned with this problem."—From Julian Bond's Foreword May/\$12.50

BASIC BOOKS, Inc.

404 Park Avenue South
New York, New York 10010

Book News in Business & Economics

Mathematical Theory of Economic Behavior

by Josef Hadar, *Case Western Reserve University*

For graduate courses in microtheory, this book follows a mathematical approach to economic behavior to demonstrate the similarities and differences between the structure of various economic models, and the methods of analyzing these models. *January (1971)*

Statistics: A Foundation for Analysis

by Ann J. Hughes and Dennis E. Grawoig, *Georgia State University*

This comprehensive discussion of the more commonly used statistical methods with business applications presents a theory of probability that is elementary and orderly, yet palatable to the non-mathematician. *March (1971)*

FORTRAN IV Computing and Applications

by Richard L. Nolan, *Graduate School of Business, Harvard University*

Designed as an introduction to computing and FORTRAN IV, this textbook studies in detail the FORTRAN IV language, its system and program flowcharting, as they relate to the computer-business environment. *April (1971)*

Introduction to Statistical Methods

by Donald L. Harnett, *Indiana University*

This book covers the basic elements of classical statistical analysis and statistical decision theory and is suitable for introductory courses in statistics. *540 pp (1970) \$10.50*

Choice Against Chance: An Introduction to Statistical Decision Theory

by John Aitchison, *University of Glasgow*

A readable, elementary introduction to decision making under uncertainty, this book stresses the relationship and correspondence of mathematical models to reality. *204 pp, 38 illus (1970) \$9.50*

Managerial Economics: Decision Making under Uncertainty for Business and Engineering

by Floyd E. Gillis, Jr., *Purdue University*

This is the first book to come to grips with the problem of applying economic theory to business decisions on the basis of conditions as they are today. *296 pp, 71 illus (1970) \$9.95*

Macroeconomic Analysis: An Introduction to Comparative Statics and Dynamics

by Thomas Dernburg and Judith Dukler Dernburg, *Oberlin College*

Intended for a second course in macroeconomics, this book studies macroeconomic problems by means of the model-building approach. *292 pp, 55 illus (1970) \$11.75*

Addison-Wesley
PUBLISHING COMPANY, INC.
Reading, Massachusetts 01867



THE SIGN OF
EXCELLENCE

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

*Have you considered these two new books
for your Fall adoptions?*

Robert L.
HEILBRONER/FORD
Graduate Faculty,
New School for Social Research

Arthur M.
FORD
Southern Illinois
University

IS ECONOMICS RELEVANT?

A Reader in Political Economics

**ROBERT
A. MUNDELL**

University of Chicago

MONETARY THEORY

**Inflation, Interest, and Growth
in the World Economy**

for approval copies write: Goodyear Publishing Company,
15115 Sunset Blvd. Pacific Palisades, California 90272

THE DEPARTMENT OF FINANCE

requires
ECONOMISTS

The Department of Finance studies the national economic situation, the government financial position and the individual and cumulative effects of various government programs. It performs a financial advisory function for the Minister who is responsible for the supervision, control and direction of matters related to the financial and economic affairs of Canada.

The Department requires economists with a relevant academic background and meaningful experience in one or more of the following areas:

Budgetary and Fiscal Planning,
Economic Analysis,
Fiscal and Economic Relations between Federal and Provincial Governments,
International Finance,
Tax Policy,
Econometrics and Statistics.

There are openings in all of these areas at different levels with salary up to \$25,000.00.

Complete details including education, experience, current salary, or salary expected should be sent directly to:

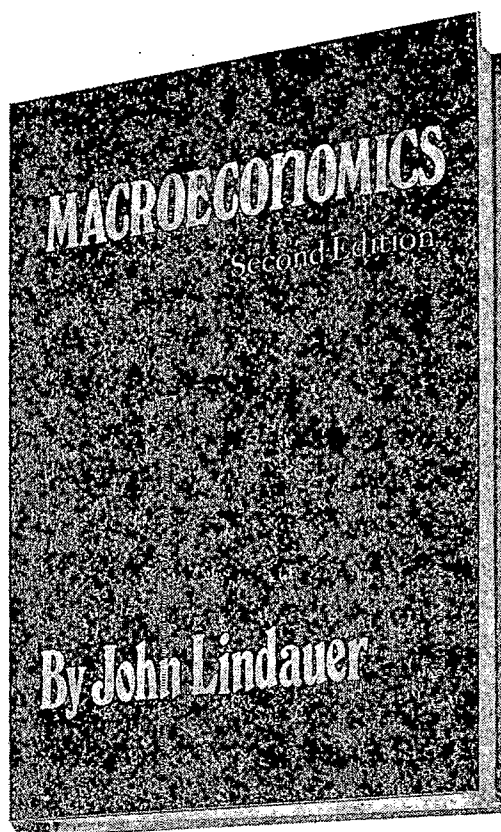


**Public
Service
of
Canada**

**MR. R. B. BONNAR, DIRECTOR
FINANCE, PERSONNEL AND ADMINISTRATION,
DEPARTMENT OF FINANCE,
OTTAWA, ONTARIO.**

Please refer to Competition 71-430-1.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



It's Concise

Lindauer did it.
He packed 20 full chapters under one cover.

Without cutting corners.

So your students get the whole essence of inflation and unemployment.

The Phillips analysis and the monetary and fiscal requirements for internal and external equilibrium.

And *more*.

The most complete discussion on aggregate supply, consumption, the macroeconomic aspects of income distribution.

In sum, the length and breadth of macroeconomics in one concise text.

Lindauer—the most coverage in the least space.

MACROECONOMICS, Second Edition
By JOHN LINDAUER, Claremont Men's
College and Claremont Graduate School
1971 In Press

For more information, contact your local Wiley representative or write Wayne Anderson, Promotion Manager.



JOHN WILEY & SONS, Inc.
605 Third Avenue, New York, N.Y. 10016
In Canada:
22 Worcester Road, Rexdale, Ontario

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Sovereignty and an Empty Purse

Banks and Politics in the Civil War

BRAY HAMMOND

In a worthy sequel to his Pulitzer Prize-winning *Banks and Politics in America*, Bray Hammond has written a spirited account of federal financing of the Civil War. Finding the ways and means to fill the Treasury's empty purse, he shows, was as important as military action in settling the war.

\$10.00

Industrialization Without National Patents

The Netherlands, 1869-1912; Switzerland, 1850-1907

ERIC SCHIFF

How vital is a national patent system for the successful industrialization of a country? Eric Schiff assesses inventive activity and technological growth in the Netherlands and Switzerland, highly sophisticated industrial countries which had no national patent protection for several decades.

\$9.50

Theory of Cost and Production Functions

RONALD W. SHEPHARD

Building upon his earlier work in the frequently cited *Cost and Production Functions* (1953), Ronald Shephard presents a detailed mathematical formulation of cost and production theory.

Princeton Studies in Mathematical Economics, 4

\$15.00

Princeton University Press

Princeton, New Jersey 08540

D. C. Heath and Company

The American Economy in Conflict: A Book of Readings

Robert B. Carson, State University of New York, Oneonta

- argues timely and controversial economic issues
- involves and stimulates beginning students

March 1971 Pb 464 pp (50815)

NEW

Contemporary Economics: Problems and Policies

*Willard W. Howard, Phoenix College
Edwin L. Dale, Jr., The New York Times*

- especially well-written
- presents difficult concepts clearly
- emphasizes current problems and policies
... just a few reasons why you should consider *Contemporary Economics* for your principles course.

Ready now Hb 544 pp (52639)

NEW

Introduction to Mathematical Economics: Matrix Algebra and Linear Economic Models

Richard H. Puckett, Board of Governors, Federal Reserve

- applies linear algebra to business and economic problems

Ready now Hb 256 pp (49783)

NEW

*For examination copies write:
Faculty Correspondence Manager MAER*

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



First Quarter Report

Problems in Political Economy: An Urban Perspective

David M. Gordon, National Bureau of Economic Research

- the liberal, radical, and conservative views presented here provide a framework for analyzing our most pressing problems
- contributors include: Walter Heller, Paul Sweezy, Milton Friedman, Elliot Liebow, Lester Thurow, Barry Bluestone, Michael Reich, Bennett Harrison, Samuel Bowles, Herbert Gintis, Anthony Downs, Jimmy Breslin, Christopher Green, Claude Brown, Bernard Frieden, John Meyer, and many others.

March 1971 Pb 544 pp
(61218)

Perspectives on Poverty and Income Distribution

James G. Scoville, University of Illinois

- readings on the economic aspects of poverty in the U.S.
- discusses income distribution, causes of poverty, possible methods of alleviating the problem

March 1971 Pb 400 pp
(60046)

Internal Labor Markets and Manpower Analysis

*Peter B. Doeringer, Harvard University
Michael J. Piore, Massachusetts Institute of Technology*

- develops an innovative theory of the internal labor market
- using this theory, analyzes the factors that influence labor economics

March 1971 Hb/Pb 288 pp
(60269)

NEW

NEW

NEW



D.C. Heath and Company

125 Spring Street
Lexington, Massachusetts 02173

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

A Recommended Selection in Economics

THE URBAN ECONOMY. By Frederick Durr, University of Delaware.

This is a discussion of the major urban economic problems facing society; a study of the economic growth and development of cities. It features the use of basic economics as a tool to understand current problems and issues. April, 1971 (paper)

ECONOMIC PROGRESS AND PROBLEMS OF LABOR, 2D ED. By Friedrich Baerwald, Fordham University

The central theme is *Manpower* in this comprehensive analysis of wage structures, full employment in the emerging white collar economy, industrial relations, and the unsolved problem of American Social Security. 1970/493 pp./\$9.25

INTERNATIONAL ECONOMICS. By David Young, University of Dundee (Scotland)

An introduction to the theory of international trade and current international monetary problems, this book deals with a theoretical analysis of international trade. Coverage includes: factor prices, customs unions theory, balance of payments adjustments, exchange rate theory, problems of disequilibrium, and current international monetary problems. 1970/246 pp./\$7.00

BASIC ECONOMICS. By Micha Gisser, The University of New Mexico; and Peter S. Barth, Ohio State University

This is a concise treatment of basic economic analysis and policies that features the inclusion of recent developments in contemporary research. Most of the chapters end with a short summary of ideas and questions that have been challenging the profession in recent years. 1970/455 pp./\$8.00

INTRODUCTION TO PRICE THEORY, 2D ED. By Micha Gisser, The University of New Mexico

The basic elements of price theory are presented by using simple tools such as tables and diagrams. Additions to this Second Edition include information on the theory of capital budgeting, general equilibrium and welfare economics, linear programming, and monopoly and imperfect competition. 1969/413 pp./\$7.95

THE ECONOMIC AND FINANCIAL SYSTEM. By Richard A. Ward, University of Southern California

This is a presentation of macroeconomic theory which emphasizes an integrated treatment of income flows and financial assets. 1970/302 pp./\$8.50

MONEY, BANKING, AND INCOME: THEORY AND POLICY. By Pham Chung, The University of New Mexico

This is an integrated treatment of the historical, institutional, and theoretical aspects of money and banking. 1970/655 pp./\$10.50

PROBABILITY AND STATISTICAL ANALYSIS. By Edgar P. Hickman and James G. Hilton, University of South Carolina

The statistical methods presented cover the most important topics including variance and multiple regression. February, 1971

Examination copies are available upon request from those in a position to consider them for adoption.



Intext Educational Publishers
(formerly International Textbook Company)
College Division
Scranton, Pennsylvania 18515

Recent and forthcoming—

The War Economy of the United States Readings in Military Industry and Economy

Seymour Melman, editor

Readings from a wide range of popular and professional literature brought together to supplement the traditional "principles of economics" and industrial management textbooks. Commonly, the principles texts do not consider—or they explain inadequately—the unique performance of military industry and economy. Here empirical economic science is used to make clear how the military-industrial firm differs from the business enterprise of industrial capitalism and to demonstrate the effects of the military economy overall on the larger economy of the United States. The selections are keyed to pertinent topics covered by conventional textbooks both in the table of contents and in the text itself.

March about 350 pages paper, about \$3.95

Problems of a Mature Economy

F. V. Meyer, D. C. Corner, and J. E. S. Parker

A study of the forces influencing the growth of individual firms, problems of industrial organization, and the allocation of resources within the British economy in the 1950s and 1960s. Part I studies management problems of the firm and the reactions of management to growth and technological change. Part II identifies problems of industrial organization—the forces of competition, monopolistic tendencies and their control, and pricing problems. Part III analyzes the distribution of the national income between the various factors of production. Part IV discusses the pattern of imports and exports and the sterling exchange rate.

February 672 pages \$16.00

Readings in the Theory of Growth

F. H. Hahn, editor

A selection of important papers on economic growth, first published in the *Review of Economic Studies*. The collection includes J. V. Newmann's "A Model of General Economic Equilibrium," the first detailed examinations of neo-classical growth models, and some of the major criticisms levelled against these. Also included are full analyses of economies with fixed coefficients and of vintage models, as well as Arrow's papers on "Learning by Doing."

April 250 pages \$12.00

ST. MARTIN'S PRESS

college department

175 fifth avenue

new york 10010

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Oxford University Press



Economic Theory and the Underdeveloped Countries

By H. MYINT, *London School of Economics*. This book, a selection of fourteen of Professor Myint's major papers from various journals and conference volumes, interprets the challenge of world poverty in terms of four themes. The first is an attempt to determine the applicability of existing economic theory to the economics of the underdeveloped countries. The second theme is concerned with the application of international trade theory to these countries. The third examines the relationships between education and economic development. And the fourth theme explores the relationship between the internal aspects, including economic planning, and the external aspects of development.

1971 336 pp. cloth \$6.95 paper \$3.95

Leading Issues in Economic Development: Studies in International Poverty

SECOND EDITION

By GERALD M. MEIER, *Stanford University*. "This is an excellent course book, indispensable for anyone interested in problems of Economic Development."—Abdul Qayum, *Portland State University*

"Still the leading volume in its field—indispensable."—Gary Gappert, *University of Wisconsin, Milwaukee*

1970 800 pp. \$9.95

Growth Theory:

AN EXPOSITION

By ROBERT M. SOLOW, *Massachusetts Institute of Technology*. "It bears the characteristic of all of Professor Solow's writing. In one word: outstanding."—Vito Tanzi, *The American University*

"A magnificent achievement."—M. Roman, *Jersey City State College*

1970 148 pp. cloth \$4.50 paper \$2.95

The Role of Science and Technology in Developing Countries

By GRAHAM JONES, *United Kingdom Institute of Statisticians*. Exploring the ways by which modern science and technology can help promote economic growth in the less developed countries, this book derives from a study sponsored by the International Council of Scientific Unions. The topics covered include science policy; the organization of research and related scientific and technological activities; their application to industry, to agriculture, and to the exploitation of natural resources; selection of appropriate technologies; and education and manpower.

Spring 1971 164 pp. cloth \$6.95 paper \$3.50

On Keynesian Economics and the Economics of Keynes:

A STUDY IN MONETARY THEORY

By AXEL LEIJONHUFVUD, *University of California, Los Angeles*. "Leijonhufvud brings to his task a splendid mastery, both of the relevant techniques and of the relevant literature. The result is an analysis which puts Keynes in a perspective radically different from that in which it has been customary to view him . . . An important book . . . A significant contribution to knowledge."—Lord Robbins, in *London, England Financial Times*

1968 446 pp. \$7.95

Society and Economic Growth:

A BEHAVIORAL PERSPECTIVE OF SOCIAL CHANGE

By JOHN H. KUNKEL, *Arizona State University*.

" . . . an extremely sensible and carefully reasoned study of the problems of bringing about economic development. The behavioral approach Kunkel uses, if in fact applied, is more likely to create real economic development than any of the other approaches that have been tried. It has a better basis in tested knowledge."—George C. Homans, *Harvard University*

1970 400 pp. \$6.50

OXFORD UNIVERSITY PRESS

200 Madison Avenue, New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Oxford University Press



The Euro-Dollar Market

By E. WAYNE GLENDENNING, *Richardson Securities of Canada*. In this study, by introducing the Euro-dollar market into relatively simple economic models, the author examines the mechanism through which it has an impact on national economies and on the international financial system, and draws conclusions about the significance of the market at the present time. 9 text-figures. \$6.75

Catholic Radicals in Brazil

By EMANUEL DE KADT, *University of Sussex*. This book traces the emergence and development of the ideas of a small but significant section of Brazil's Catholic intelligentsia which attempted to help the underprivileged to find solutions to their problems. It focuses on the Movement for Basic Education, its field of operation, and its main tools. (*For the Royal Institute of International Affairs.*) \$9.75

Railways and Economic Growth in England and Wales 1840-1870

By G. R. HAWKE, *University of Wellington, New Zealand*. In this investigation of a long neglected topic, the author assesses the role of the railway in the economic growth of England and Wales. The reduction of transport costs effected by the railway constitutes a major part of the study. The second part of the study is a search for dynamic links between railways and growth in other industries. 48 figures, 63 tables. \$19.25

Industrial Conflict in Malaya

PRELUDE TO THE COMMUNIST REVOLT OF 1948

By M. R. STENSON, *University of Auckland, New Zealand*. In this study Dr. Stenson analyzes the internal dynamics of industrial conflict within a rapidly changing colonial society. He discusses the changes brought about by the communist revolt and its influence upon the growth and nature of unionism. The revolt of 1948 is also shown as a result of World War II having increased the differences between the labor force and the employers. A postscript surveys the subsequent period up to 1969. \$8.00

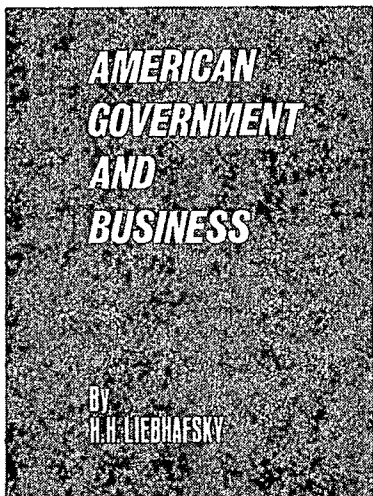
OXFORD UNIVERSITY PRESS

200 Madison Avenue, New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

LIEBHAFSKY

= The most inclusive text
+ The most up-to-date text



Some economics/government texts hem readers into coverage that's too narrow. Not Liebhafsky.

First of all, he shows government's relations with the private sector with an interdisciplinary approach. Integrating the legal with the political, the philosophical with the economic.

Secondly, *working* official papers and judicial decisions are blended into the text so students see the connections between theory and actuality.

Which means they won't need supplementary readers.

And the documentation is both timely and relevant. And varied.

In fact, Liebhafsky's broader and more current treatment includes *all* the issues. Consumer protection; pollution control; recent merger decisions; antitrust decrees and constitutional law. Even truth-in-lending, auto safety, econometric models in rate cases, and the work of the Federal Reserve Board.

AMERICAN GOVERNMENT AND BUSINESS

By H. H. Liebhafsky, *University of Texas*
1971 (approx.) 640 pages \$11.95

AND, FOR MORE ADVANCED COURSES:

Asch explores *theory*—game theory, oligopoly theory—along with collusion approaches, and a consideration of classical approaches.

Yet he rigorously examines conceptual and practical *problems* in measuring competition, antitrust, and monopoly.

With thorough discussions of policy areas in later chapters to relate the theoretical to the empirical materials covered.

ECONOMIC THEORY AND THE ANTITRUST DILEMMA

By Peter Asch, *Rutgers University*
1970 414 pages \$9.95

Kahn goes beyond just counting off government regulation of public utilities—proper rate structures, meting out common and joint costs, rate-time ratios, accommodation of competition, quality of service, rate base, rate of return and cream skimming.

He *combines* his appraisal with the relevant, basic principles of micro-economic theory and shows how they apply to these major problems.

In both books, *economic efficiency* is the core of the analysis. Volume II, just as detailed and easy to take as Volume I, applies living institutional issues to regulatory problems.

THE ECONOMICS OF REGULATION: Principles and Institutions

By ALFRED E. KAHN, *Cornell University*

Volume I: Economic Principles
1970 212 pages \$9.95

Volume II: Institutional Issues
1971 (approx.) 368 pages \$10.95

For further information, contact your local Wiley representative, or write Wayne Anderson, Promotion Manager



JOHN WILEY & SONS, Inc., 605 Third Avenue, New York, N.Y. 10016
In Canada: 22 Worcester Road, Rexdale, Ontario

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Outstanding Texts!

ASCENT TO AFFLUENCE: A HISTORY OF AMERICAN ECONOMIC DEVELOPMENT

by Charles H. Hession and Hyman Sardy, both of Brooklyn College—a unique text that integrates the theories of economic growth with a detailed analysis of the American economic experience. 1969, 6 x 9, 896 pages.

FUNDAMENTAL STATISTICS FOR BUSINESS AND ECONOMICS, Third Edition

by John Neter, University of Minnesota; William Wasserman, Syracuse University—presents a modern approach to statistics by emphasizing the application of statistical techniques as tools for decision-making. 6 x 9, 738 pages.

SELF-CORRECTING PROBLEMS IN STATISTICS

by George A. Whitmore, McGill University; John Neter, University of Minnesota; William Wasserman, Syracuse University—a unique teaching program—36 sections provide the student with the "basics" of the topics in statistics. 1970, 8 1/2 x 11, Paperbound, 342 pages.

QUANTITATIVE INTERNATIONAL ECONOMICS

by Edward E. Leamer, Harvard University; Robert M. Stern, University of Michigan—bridges the gap between pure theory and policy and brings together all the important econometric studies of international economics. 1970, 6 x 9, 209 pages.

MATHEMATICAL MODELS IN MICROECONOMICS

by Coldwell Daniel, III, Memphis State University—a clear, well-organized development of the basic models in microeconomic analysis. 1970, 6 x 9, 228 pages.

THE CONTRASTING ECONOMIES: A STUDY OF MODERN ECONOMIC SYSTEMS, Second Edition

by Lynn Turgeon, Hofstra University—presents a thorough functional treatment of the economies of capitalist and noncapitalist systems. 1969, 5 3/8 x 8 3/8, Paperbound, 445 pages.

Allyn and Bacon, Inc.

College Division, Dept. 893
470 Atlantic Avenue, Boston MA 02210

FINANCIAL INSTITUTIONS AND MARKETS

Murray E. Polakoff, New York University, and Others

Twenty-eight distinguished specialists contribute original chapters in this significant text which provides an integrated examination of financial intermediaries.

640 pages / 1970 / \$12.95

MACROECONOMICS: Selected Readings

Edited by Walter L. Johnson and David R. Kemerschen,
both of the University of Missouri

520 pages / paper / 1970 / \$6.50

PRICE THEORY IN ACTION: A Book of Readings, Second Edition

Edited by Donald Stevenson Watson,
George Washington University

400 pages / paper / 1969 / \$5.50

PRICE THEORY AND ITS USES, Second Edition

Donald Stevenson Watson, George Washington University

443 pages / 1968 / \$8.75

**DOMESTIC TRANSPORTATION: PRACTICE, THEORY, AND POLICY,
Second Edition**

Roy J. Sampson, University of Oregon and Martin T. Farris,
Arizona State University

525 pages / March, 1971

ECONOMIC ESSENTIALS: A Core Approach

Roger C. Van Tassel, Clark University

400 pages / paper / 1969 / \$5.50

PERSPECTIVES OF ECONOMIC DEVELOPMENT

Salvatore Schiavo-Campo, University of Massachusetts, Boston, and
Hans W. Singer, University of Sussex

350 pages / paper / 1970 / \$5.95

DECISION MAKING FOR ECONOMIC DEVELOPMENT: Text and Cases

Gustav F. Papanek, Daniel M. Schydrowsky, and Joseph J. Stern,
all of Harvard University

172 pages / paper / 1970 / \$4.95

Houghton Mifflin

Boston 02107 / Atlanta 30324 / Dallas 75235 / Geneva, Ill. 60134
New York 10036 / Palo Alto 94304

ANNOUNCING



COMPARATIVE ECONOMIC SYSTEMS

Martin C. Schnitzer
Virginia Polytechnic Institute and State University
James W. Nordyke
New Mexico State University



DECISIONS UNDER UNCERTAINTY

A. N. Halter
Oregon State University
G. W. Dean
University of California at Davis



BUSINESS AND ECONOMIC STATISTICS, *Second Edition*

John R. Stockton
University of Texas at Austin
Charles T. Clark
University of Texas at Austin



FRONTIERS OF FINANCIAL MANAGEMENT, *Sixth Edition*

William J. Serraino
Miami University, Oxford
Surendra S. Singhvi
Miami University, Oxford
Robert M. Soldofsky
The University of Iowa



STRATEGY, POLICY, AND CENTRAL MANAGEMENT

William H. Newman
Columbia University
James P. Logan
University of Arizona



SOUTH-WESTERN PUBLISHING CO.

(Specialists in Business and Economic Education)

Cincinnati, OH 45227 • Chicago, IL 60644 • Dallas, TX 75229
New Rochelle, NY 10802 • Burlingame, CA 94010 • Brighton, England

Random House—Alfred A. Knopf

Teach your students to analyze, not memorize.

Too many texts present economics as a kind of jigsaw puzzle, broken up into pieces. But they don't finish the puzzle.

They don't explain to the student how the elements fit together. He leaves the class with a head full of formulas he can rattle off on command, but no real understanding.

Miernyk puts it all together.

Approaching economics as a system of interdependent elements, he shows how each is part of the whole.

Miernyk provides the student with the tools for analyzing the workings of the economic system and its problems.

ECONOMICS

William H. Miernyk, West Virginia University

Random House, 1970

672 pages; \$9.95

An Instructor's Manual and a Programmed Review Guide and Workbook are also available.

Locked in by texts that teach yesterday's economics?

Break away with the reader that plunges the student into the economics of today—

ECONOMICS Mainstream Readings and Radical Critiques

Edited by David Mermelstein, The Polytechnic Institute of Brooklyn, with a special introduction by Robert Lekachman

Professor Mermelstein presents a challenging confrontation between orthodox and radical economists by pairing their articles, showing the student the controversy surrounding many current economic problems. This dynamic anthology provides a unique and refreshing counterpoint to whatever basic text the instructor chooses.

Random House, 1970

640 pages; \$5.95 paperbound

New—

A PRIMER ON GOVERNMENT SPENDING Second Edition

Robert L. Heilbroner, New School for Social Research, and Peter L. Bernstein, New School for Social Research and Bernstein-Macauley, Inc.

Random House, 1970

120 pages; \$1.95 paperbound

COMPARATIVE ECONOMIC SYSTEMS A Reader Second Edition

Marshall I. Goldman, Wellesley College

The second edition includes new sections on capitalism in Japan and France, material on the nature and goals of economic systems, a discussion of contemporary theory and critiques of American capitalism. The selections and introductions carry the student beyond "isms" and into more critical examinations of the methods by which men have tried to solve the problem of scarcity.

Random House, 1971

456 pages; \$6.95 paperbound



RANDOM HOUSE

The College Department

201 East 50th Street, New York 10022

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Random House—Alfred A. Knopf

Just published—

MACROECONOMICS

Norman F. Keiser, San Jose State College

Presenting the fundamentals of macroeconomic theory on the intermediate level, Professor Keiser explains the relevance of macroeconomics to contemporary society by emphasizing empirical studies and policy considerations. Breaking away from the usual pattern, this text devotes less space to national income analysis and spends proportionately more time on investment, fiscal policy, aggregate supply, growth theory, and policy decisions.

Instructor's Manual

Random House, March 1971

512 pages; \$9.95

To accompany MACROECONOMICS or any standard text:

MACROECONOMICS Problems and a Survey of Theory

Norman F. Keiser

An outline of macroeconomic theory is a unique feature of this workbook. The problems and exercises are designed so that the student plays a major role in building models for income expenditure, real goods, and the monetary sector. The inclusion of the problem and exercise solutions enables the student to test his understanding of the material.

Random House, 1971

256 pages; \$4.95 paperbound

DIMENSIONS OF MACROECONOMICS A Book of Readings

Edited by Sid Mittra, Oakland University

Forty-four major articles from a wide variety of national and international sources. Covering each area of macroeconomics, the readings concentrate on the use of macroeconomic tools to achieve long term economic goals and to solve such problems as unemployment, poverty, inflation, and low income. An unusual feature of this reader is its inclusion of additional influential articles in a summarized form.

Random House, April 1971

384 pages; \$6.00 paperbound

MONETARY ECONOMICS Controversies in Theory and Policy

Edited by Jonas Prager, New York University

This reader examines a limited number of current issues in depth. Dr. Prager focuses on two major controversies: the Keynesian version of the monetary mechanism versus the neo-quantity theory, and the empirical verification of monetary theory.

Random House, 1971

448 pages; \$6.00 paperbound



RANDOM HOUSE

The College Department

201 East 50th Street, New York 10022

Heath Lexington Books

Forthcoming Titles

Internal Labor Markets and Manpower Analysis.

Peter B. Doeringer, Harvard University, Michael J. Piore, Massachusetts Institute of Technology. Develops an innovative theory of the internal labor market which suggests that the enterprise is an important and somewhat insulated labor market institute. Analyzes factors that influence labor economics such as job training, automation and government manpower policy theory and policy application of labor economics.

ca. 224 pp. March \$7.50
paper \$3.95

Industrial Location in the United States.

James C. Burrows, Charles E. Metcalf, John B. Kaler, Charles River Associates, Inc. Makes a significant contribution to the literature of regional development by identifying the factors that cause changes in an area's industry structure. Develops models for estimating an area's future industrial mix.

200 pp. March \$15.00

Cobalt: An Industry Study

James C. Burrows, Charles River Associates, Inc. The first in a series of full-scale studies of the minerals industries, this book combines industry and econometric analysis to provide an understanding of the supply and

demand responses of the cobalt industry and its structure. Uses econometric evidence to test different models of market behavior.
256 pp. March \$25.00

The Role of Transportation in Regional Development.

Charles River Associates, Inc. An analysis of the complex interrelationship between transportation and regional economic development. Discusses the active role of transportation in terms of regional economics, international trade theory and location theory.
112 pp. May ca. \$15.00

The Use of Econometric Models by Federal Regulatory Agencies.

Joe L. Steele, Texas Christian University. Estimates the potential impact of econometric models on rate setting by Federal regulatory agencies.
ca. 120 pp. May ca. \$12.50

Available Titles

Predictability of Stock Market Prices.

Clive W. J. Granger, University of Nottingham, Oskar Morgenstern, New York University.
336 pp. \$15.00

Speculation, Hedging and Commodity Price Forecasts.

Walter C. Labys, United Nations Conference on Trade and Development, Clive W. J. Granger, University of Nottingham.
352 pp. \$15.00

The Demand for Travel: Theory and Measurement.

Edited by Richard E. Quandt, Princeton University.
320 pp. \$15.00

International Travel-International Trade.

H. Peter Gray, Rutgers University. 286 pp. \$12.50

An Analysis of the United States Oil Import Quota.

James C. Burrows, Thomas A. Domencich, Charles River Associates, Inc.
280 pp. \$19.50

Simulating Regional Economic Development.

William H. Miernyk et al., University of West Virginia.
365 pp. \$15.00

A Short Run Forecasting Model of the United States Economy.

Ray C. Fair, Princeton University.
274 pp. \$12.50

Heath LEXINGTON Books

D.C. Heath and Company
125 Spring Street
Lexington, Massachusetts
02173

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

SURVEYS OF AFRICAN ECONOMIES

Volume 3: Dahomey, Ivory Coast, Mauritania, Niger, Senegal, Togo, and Upper Volta

PP. xxxviii + 786

Maps, Index, Statistical Tables

This volume, covering the seven members of the West African Monetary Union, is the third in the series of books describing the economies of selected African countries. Each volume surveys the monetary, fiscal, exchange control, and trading systems that are common to the groups of selected countries and presents detailed information for each country on natural resources, production, development planning and progress, budgets and taxation, money and banking, and foreign trade, aid, and payments.

Previous volumes in the series are

Volume 1: Cameroon, the Central African Republic, Chad, Congo (Brazzaville), and Gabon

Volume 2: Kenya, Tanzania, Uganda, and Somalia

Price: \$5.00 a volume (\$2.50 to university libraries, faculty members, and students).

Address correspondence to

The Secretary

INTERNATIONAL MONETARY FUND

19th and H Streets, N.W.

Washington, D.C. 20431 U.S.A.

YOUR BOOKS AND JOURNALS NEEDED IN ASIA

Books for Asian Students is seeking donations of books published since 1958 and professional and technical journals in 10 year runs commencing 1946 or later for distribution to institutions and individuals in Asia.

Members of professional associations often have materials highly relevant and applicable to present day Asian needs. You can "recycle" this resource. By donating your books, you put them back to work in the hands of people much like yourself: people seeking knowledge and skills, charged with responsibility, and concerned about their fellow man and about the future.

Books that go abroad assist the development process; they also foster the appreciation of culture, values, and beliefs which is essential for true international understanding.

For details of what is needed and where to send it write:

Carlton Lowenberg, Director

Books for Asian Students

451 Sixth Street

San Francisco, California 94103

[Telephone: (415) 982-4640]



THE UNIVERSITY OF CHICAGO PRESS

5801 South Ellis Avenue Chicago, Illinois 60637

FOUNDATIONS, PRIVATE GIVING, AND PUBLIC POLICY
Report and Recommendations of the Commission on Foundations and Private Philanthropy

Foundations—public drain or private gift? Answering the current onslaught of criticism, which includes Congressional charges that foundations are mismanaging their funds and depriving the government of badly-needed taxable dollars, this unique report makes public original research on foundation activity. Composed of sixteen prominent educators, lawyers, and businessmen, the blue-ribbon commission concludes that "the public interest is best served through a strong dual system of private giving and government funding."

The commission bases its judgment on an in-depth survey of over 200 representative American foundations. Its report provides the most up-to-date and detailed information presently available on foundation funding, grant-making patterns, and investment histories; its recommendations to both government and foundation executives plot guidelines for the future development of private philanthropy.

1971 5½ × 8½ 312 pages Cloth \$12.50

RISK, UNCERTAINTY, AND PROFIT

Frank H. Knight

"This volume is clearly one of the half-dozen classics in economic theory to appear in this country in the first half-century—to be compared to Fisher's *Theory of Interest* or Mitchell's *Business Cycles*. It has become a standard piece in the literature of general theory as well as of uncertainty and economic change." (George J. Stigler)

"It is not only one of the best short statements of the neoclassical theory of value (with a distinctly 'Austrian' twist) but it made an important contribution to the theory of profit." (Kenneth E. Boulding)

"This is certainly a classic. Time has not made it obsolete." (Paul A. Samuelson)

Risk, Uncertainty, and Profit was originally published in 1921.

P 396 1971 5¼ × 8 394 pages Paper \$3.95

THE FISCAL REVOLUTION IN AMERICA

Herbert Stein, President's Council of Economic Advisers

Between 1931 and 1962, the principle of balancing the budget gave way to the principle of managing government expenditures and taxes to achieve prosperity, with some attention also to the supply of money and credit. The history of the unfinished revolution helps to explain the current status of economic policy in America and the prospects of the American economy. Mr. Stein tells the story in terms of its leading participants and their thinking and action at critical points in the past forty years.

"Sober, well-written, beautifully documented . . . a combined exercise in economic and intellectual history, a kind of book that is all too rare." (Kenneth Boulding)

"Dr. Stein is one of that larger than life band of American economists who combine first order academic excellence with its application to business thinking and government policy." (*The Economist*)

Cloth edition published 1969 (\$10.00).

1971 6 × 9 544 pages Paper \$4.50

ANNOUNCING . . .

a new series . . .

of inexpensive, paperbound essays, designed to provide concise and authoritative studies of the men who have made the greatest contributions to modern economic thought.

Columbia Essays on Great Economists

Donald Dewey, General Editor

The first three volumes to be published:

1. Karl Marx by Murray Wolfson
2. Edward H. Chamberlin by Romney Robinson
3. Eugen von Böhm-Bawerk by Robert Kuenne

\$1.45 each

Columbia University Press

Address for orders: 136 So. Broadway, Irvington, N.Y. 10533

In Canada: McGill-Queen's University Press, 3458 Redpath St., Montreal 109

THE ECONOMIC INSTITUTE Announces its 14th Summer Program of

Foreign Students

to be held at the University of Colorado, June 16–August 18, 1971

The Institute is designed for students planning to undertake graduate studies in the United States during the next academic year. Intensive course work is offered in each of the following areas, adjusted to individual student needs:

ECONOMIC THEORY ■

ENGLISH ■

MATHEMATICS ■

STATISTICS ■

The primary purpose is to help students review and strengthen their academic preparation. However, course work is supplemented by special lectures, seminars, field trips and social and recreational activities designed to also acquaint them with the United States economy, campus and community life.

REQUIREMENT: All students attending the Economics Institute must provide proof of admission to an acceptable graduate school program in the United States for the following academic year.

INFORMATION: For further information and application materials write to: DIRECTOR, ECONOMICS INSTITUTE, UNIVERSITY OF COLORADO, BOULDER, COLORADO 80302, U.S.A.

Program sponsored by the



American Economic Association

in cooperation with Admitting Universities, the Institute of International Education and other Foreign Student Sponsoring Agencies

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

THE ANALYSIS OF ECONOMICS

MACROECONOMICS BEFORE KEYNES

By James L. Cochrane, The University of South Carolina

The idea is held by many beginning students that economics did not exist before 1936. This selective history of macroeconomics demonstrates the remarkable homogeneity of macroeconomic theory, by examining the developments and continuing influence of work done before Keynes. Each of four pre-Keynesian schools is discussed in an individual chapter: physiocratic, classical, Marxian, and neoclassical, with a simple aggregate model as the core of each discussion. A degree of continuity and symmetry among the chapters is maintained throughout, even though a different tool of economic analysis is stressed in the construction of each model.

©1970; 109 pages, softbound \$2.75

THE AGE OF THE ECONOMIST

By Daniel R. Fusfeld, The University of Michigan

A lucid, literate exposition of the major trends in economic thought, *The Age of the Economist* presents the ideas of economists in relation to the political and philosophical ideas and economic circumstances of their ages. The author's application of the "sociology of knowledge" approach emphasizes the significance of economics for public policy issues, ideological disputes, and social philosophies. The book is a useful and interesting supplement to standard economics texts where instructors wish to include material on the development of economics in its intellectual and historical setting.

©1966, 147 pages, hardbound \$5.50 and softbound \$3.50

ECONOMICS: Analysis and Issues

By Thomas Sowell, University of California, Los Angeles

This concise text views economics as a *system* of analysis, a framework of ideas built upon one another, and solidly linked together to bring out the real-world meaning of economic facts and theory and their application to concrete problems and issues.

The emphasis throughout is on systematic analysis, rather than on either factual information or theoretical elegance. Discussions of controversial policy issues are strongly presented as opposing viewpoints, which confront the student with the hard choices of the real world instead of neatly packaged "solutions."

The analytical substance of the book is presented in the first three parts, beginning with the basic problems of scarcity, allocation, and substitution, and proceeding to macroeconomics. This material is illustrated with numerous examples and problems. By contrast, Part Four is arranged by topic, so that students can tackle some problems whole, instead of seeing only those aspects of the problem which happen to fit a particular chapter's development of analytical tools.

©1971, approx. 400 pages, hardbound \$7.95 (tent.)

ECONOMICS, Readings in Analysis and Policy

Edited by Dennis R. Starleaf, Iowa State University

The selections in this book have more analytical content than those contained in most readings books intended for use in the principles of economics course. The selections were written by economists for consumption by the intelligent layman. The readings function around three objectives: (1) to enrich and reinforce understanding of fundamental and important analytical concepts, such as opportunity costs, supply and demand, marginal costs, marginal benefits, and circular flow; (2) to illustrate the application of analytical concepts and techniques to modern economic problems; and (3) to present thoughtful material on current economic problems and policy issues in greater depth than standard principles texts. The arrangement and approach closely parallels that of the text described above.

©1969, 525 pages, softbound \$4.95



For further information

write to Joel Baron, Advertising Department

Scott, Foresman College Division

1900 East Lake Avenue Glenview, Illinois 60025



How do you read a **microbook** volume?

Comfortably.

As a book should be read. In an easy chair. At a table or desk. Almost anyplace that suits you.

The reason? A new lap reader. The Microbook 710. It has been specially designed for reading Microbook libraries, such as the 20,000-volume Library of American Civilization. It takes a 3" x 5" Microbook fiche with up to 1,000 pages and clearly projects a page at a time in a 7" x 10" size. You can read in normal room light. Finding pages, or scanning through a whole book, is easier than with a regular book.

Library Resources also offers a table reader with a bright 8½" x 12" screen. And a reader-printer is on the way.

The Library of American Civilization is the first of many libraries to be offered in Microbook form. For storage, retrieval, check-out, and reading convenience, there is one book per fiche for most single-volume works.

The Library includes the most significant 20,000 volumes of retrospective source materials on Americana. It is fully referenced and cross-indexed with author, title, and subject catalogs plus a 2,000-page Biblioguide™ Index, or topical index. Fifteen sets of these come with each library. So, faculty members may have their personal set of the catalogs and the Biblioguide Index.

The Library of American Civilization comes complete and is ready to use the day you receive it. That's a comfort, too!

For complete information, and a copy of our new 20-page booklet, "The Microbook Library Series," please write to the address below.

Library Resources, Inc.

an ENCYCLOPAEDIA BRITANNICA Company,
Dept. C-1, 301 East Erie Street, Chicago, Illinois 60611



Thomas D. Crocker,

University of California, Riverside

A. J. Rogers III,

University of Wisconsin, Milwaukee

ENVIRONMENTAL ECONOMICS

January 1971—160 pages \$2.95

Samuel A. Morley,

University of Wisconsin

THE ECONOMICS OF INFLATION

Summer 1971 128 pages \$2.95 tent.

Recently Published:

Richard H. Leftwich,

Oklahoma State University

**PRICE SYSTEM &
RESOURCE ALLOCATION,
4th Edition**

1970 416 pages \$9.50

Heinz Kohler,

Amherst College

**ECONOMICS: The Science
of Scarcity**

1970 748 pages \$10.50

Forthcoming:

**Eugene F. Brigham
and James Pappas,**

University of Wisconsin

MANAGERIAL ECONOMICS

Jan. 1972 512 pages \$9.95



THE DRYDEN PRESS, INC.

901 North Elm
Hinsdale, Illinois 60521

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Well-Known Harper & Row Texts...

Economics, Second Edition

RICHARD G. LIPSEY, Queen's University, Canada, and
PETER O. STEINER, The University of Michigan

1969. 845 pp.; \$10.95.

To accompany Lipsey and Steiner's Economics...

Instructor's Manual.

Study Guide by DOROTHY FORBUSH and DASCOMB R. FORBUSH, Clarkson College of Technology (\$3.50).

The Economics of Money and Banking, Fifth Edition

LESTER V. CHANDLER, Atlanta University

1969. 562 pp.; \$11.95.

To accompany Chandler's The Economics of Money and Banking, 5th Ed. ...

Instructor's Manual by HAROLD R. WILLIAMS, Harvard University, and HENRY W. WOUTENBERG, Kent State University.

Study Guide by HAROLD R. WILLIAMS and HENRY W. WOUTENBERG (\$3.95).

Money, Banking, and Monetary Policy: Readings in Domestic and International Policy edited by HAROLD R. WILLIAMS and HENRY W. WOUTENBERG (724 pp.; \$8.95; paper).

Mathematical Analysis: Business and Economic Applications

JEAN E. DRAPER, University of Arizona, and

JANE S. KLINGMAN, Xavier University.

1967. 568 pp.; \$13.95.

To accompany Draper and Klingman's Mathematical Analysis ...

Solutions Manual.

Principles of Economics

DANIEL B. SUITS, University of California, Santa Cruz

1970. 542 pp.; \$9.95.

To accompany Suits' Principles of Economics ...

Instructor's Manual.

Student Guide (\$3.95).



1817

Send for our 1971 catalog /#CT 38

HARPER & ROW, PUBLISHERS, INC.

49 East 33d Street, New York 10016

New additions... to Harper & Row's list...

Introductory Economics

RONALD A. WYKSTRA, Colorado State University

Presenting economic theory from a socioeconomic viewpoint, this basic full-course text covers not only standard topics but also such topics as welfare and the public sector and environmental economics. Uses established economic theory, with macroeconomics preceding microeconomics. *Instructor's Manual* by Ronald A. Wykstra and Jack Whippen. *Lecture Notes* by Ronald A. Wykstra. March, 1971. *Tentative*: 750 pp.; \$10.95.

To accompany Wykstra's Introductory Economics . . .

Study Guide by JOHN R. McKEAN, Colorado State University, and
RONALD A. WYKSTRA (*Tentative*: 200 pp.; \$3.95).

Readings in Introductory Economics edited by JOHN R. McKEAN
and RONALD A. WYKSTRA (*Tentative*: 320 pp.; \$4.50; *paper*).

The Reinterpretation of American Economic History

edited by ROBERT W. FOGEL, University of Chicago and University of
Rochester, and STANLEY L. ENGERMAN, University of Rochester
April, 1971. Tentative: 480 pp.; \$12.95.

European Economic History: Tenth Century to the Present

ELIAS H. TUMA, University of California, Davis
April, 1971. 384 pp.; \$12.95.

International Trade: An Introduction to Method and Theory

A. MYRICK FREEMAN III, Bowdoin College
February, 1971. 231 pp.; \$7.95.

The Economics of Public Issues

DOUGLASS C. NORTH and ROGER LeROY MILLER,
University of Washington
April, 1971. Tentative: 192 pp.; \$2.95. *Paper.*



1817

Send for our 1971 catalog/#CT 38

HARPER & ROW, PUBLISHERS, INC.

49 East 33d Street, New York 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

The American Economic Review

Published by the American Economic Association

1970

Volume 60, Number 1

January

Published by the American Economic Association, 1015 15th Street, N.W., Washington, D.C. 20004

Subscription price, \$10.00 per volume (1970)

Single copies, \$3.00 each (1970)

1970

AMERICAN ECONOMIC ASSOCIATION
Organized at Saratoga, New York, September 9, 1885

PAST OFFICERS

Presidents

1. FRANCIS A. WALKER,* M.I.T., 1886-92
2. CHARLES F. DUNBAR,* Harvard, 1893
3. JOHN B. CLARK,* Columbia, 1894-95
4. HENRY C. ADAMS,* Michigan, 1896-97
5. ARTHUR T. HADLEY,* Yale, 1898-99
6. RICHARD T. ELY,* Wisconsin, 1900-01
7. EDWIN R. A. SELIGMAN,* Columbia, 1902-03
8. FRANK W. TAUSSIG,* Harvard, 1904-05
9. JEREMIAH W. JENKS,* Cornell, 1906-07
10. SIMON N. PATTEN,* Pennsylvania, 1908
11. DAVIS R. DEWEY,* M.I.T., 1909
12. EDMUND J. JAMES,* Illinois, 1910
13. HENRY W. FARNAM,* Yale, 1911
14. FRANK A. FETTER,* Princeton, 1912
15. DAVID KINLEY,* Illinois, 1913
16. JOHN H. GRAY,* Minnesota, 1914
17. WALTER F. WILLCOX,* Cornell, 1915
18. THOMAS N. CARVER,* Harvard, 1916
19. JOHN R. COMMONS,* Wisconsin, 1917
20. IRVING FISHER,* Yale, 1918
21. HENRY B. GARDNER,* Brown, 1919
22. HERBERT J. DAVENPORT,* Cornell, 1920
23. JACOB H. HOLLANDER,* Johns Hopkins, 1921
24. HENRY R. SEAGER,* Columbia, 1922
25. CARL C. PLEHN,* California, 1923
26. WESLEY C. MITCHELL,* Columbia, 1924
27. ALLYN A. YOUNG,* Harvard, 1925
28. EDWIN W. KEMMERER,* Princeton, 1926
29. THOMAS S. ADAMS,* Yale, 1927
30. FRED M. TAYLOR,* Michigan, 1928
31. EDWIN F. GAY,* Harvard, 1929
32. MATTHEW B. HAMMOND,* Ohio State, 1930
33. ERNEST L. BOGART,* Illinois, 1931
34. GEORGE E. BARNETT,* Johns Hopkins, 1932
35. WILLIAM Z. RIPLEY,* Harvard, 1933
36. HARRY A. MILLIS,* Chicago, 1934
37. JOHN M. CLARK,* Columbia, 1935
38. ALVIN S. JOHNSON, New School, 1936
39. OLIVER M. W. SPRAGUE,* Harvard, 1937
40. ALVIN H. HANSEN, Harvard, 1938
41. JACOB VINER,* Chicago, 1939
42. FREDERICK C. MILLS,* Columbia, 1940
43. SUMNER H. SLICHTER,* Harvard, 1941
44. EDWIN G. NOURSE, Brookings, 1942
45. ALBERT B. WOLFE,* Ohio State, 1943
46. JOSEPH S. DAVIS, Stanford, 1944
47. I. L. SHARFMAN,* Michigan, 1945
48. E. A. GOLDENWEISER,* Institute for Advanced Study, 1946
49. PAUL H. DOUGLAS, Chicago, 1947
50. JOSEPH A. SCHUMPETER,* Harvard, 1948

51. HOWARD S. ELLIS, California, 1949
52. FRANK H. KNIGHT, Chicago, 1950
53. JOHN H. WILLIAMS, Harvard, 1951
54. HAROLD A. INNIS,* Toronto, 1952
55. CALVIN B. HOOVER, Duke, 1953
56. SIMON KUZNETS, Pennsylvania, 1954
57. JOHN D. BLACK,* Harvard, 1955
58. EDWIN E. WITTE,* Wisconsin, 1956
59. MORRIS A. COPELAND, Cornell, 1957
60. GEORGE W. STOCKING, Vanderbilt, 1958
61. ARTHUR F. BURNS, Columbia, 1959
62. THEODORE W. SCHULTZ, Chicago, 1960
63. PAUL A. SAMUELSON, M.I.T., 1961
64. EDWARD S. MASON, Harvard, 1962
65. GOTTFRIED HABERLER, Harvard, 1963
66. GEORGE J. STIGLER, Chicago, 1964
67. JOSEPH J. SPENGLER, Duke, 1965
68. FRITZ MACHLUP, Princeton, 1966
69. MILTON FRIEDMAN, Chicago, 1967
70. KENNETH E. BOULDING, Colorado, 1968
71. WILLIAM J. FELLNER, Yale, 1969
72. WASSILY LEONTIEF, Harvard, 1970

Secretaries

- RICHARD T. ELY,* 1886-92
EDWARD A. ROSS,* 1893
JEREMIAH W. JENKS,* 1894-96
WALTER F. WILLCOX,* 1897-99

Treasurers

- EDWIN R. A. SELIGMAN,* 1886-90
FREDERICK B. HAWLEY,* 1891-95
CHARLES H. HULL,* 1896-99

Secretary-Treasurers

- CHARLES H. HULL,* 1900
FRANK A. FETTER,* 1901-06
WINTHROP M. DANIELS,* 1907-08
THOMAS N. CARVER,* 1909-13
ALLYN A. YOUNG,* 1914-20
RAY B. WESTERFIELD,* 1921-25
FREDERICK S. DEIBLER,* 1925-35
JAMES WASHINGTON BELL,* 1936-61
HAROLD F. WILLIAMSON, 1962-70

Editors

- DAVIS R. DEWEY,* 1911-40
PAUL T. HOMAN,* 1941-51
BERNARD F. HALEY, 1952-62
JOHN G. GURLEY, 1962-68

* Deceased.

OFFICERS FOR 1971

President

JAMES TOBIN, Yale University

President-elect

JOHN KENNETH GALBRAITH, Harvard University

Vice-Presidents

FRANCO MODIGLIANI, Massachusetts Institute of Technology
JAMES M. BUCHANAN, Virginia Polytechnic Institute

Secretary-Treasurer and Editor of Proceedings
RENDIGS FELS, Vanderbilt University

Managing Editor of the American Economic Review
GEORGE H. BORTS, Brown University

Editor of the Journal of Economic Literature

MARK PERLMAN, University of Pittsburgh

Elected Members of the Executive Committee

MARY JEAN BOWMAN, University of Chicago
CHARLES L. SCHULTZE, Brookings Institution
ROBERT DORFMAN, Harvard University
ARNOLD C. HARBERGER, University of Chicago
JOHN R. MEYER, Yale University
ROBERT EISNER, Northwestern University

Ex Officio Members

WILLIAM J. FELLNER, Yale University
WASSILY LEONTIEF, Harvard University

Counsel

SEYMOUR J. RUBIN, Washington, D.C.

RICHARD T. ELY LECTURE

The Keynesian Revolution and the Monetarist Counter-Revolution

By HARRY G. JOHNSON

The London School of Economics and Political Science and The University of Chicago

When James Tobin and I agreed on the subject of this lecture last spring, it appeared to be a highly topical subject that would command widespread interest among the membership of this Association. Unfortunately, as so often happens with forward planning for academic purposes, others have also been alert to topicality, and have undermined our forward planning by getting in earlier with their version of the theme. Thus Milton Friedman himself gave a widely publicized lecture on "The Counter-Revolution in Monetary Theory" last September in London, which lecture has recently been published by the Institute of Economic Affairs, [4]; Karl Brunner has recently circulated a typically scholarly paper on "The 'Monetarist Revolution' In Monetary Theory" [1]; and undoubtedly many others have been writing and publishing on the same subject. My treatment of this beginning-to-be-well-worn theme today will, I hope, still retain some novelty, inasmuch as I shall be primarily concerned, not with the scientific issues in dispute in the monetarist counter-revolution against the Keynesian revolution, but with the social and intellectual conditions that make a revolution or counter-revolution possible in our profession. This lecture is therefore an excursion—amateurish, I must confess—into the economics and sociology of intellectual change.

As is well known from the field of economic history, the concept of revolution is

difficult to transfer from its origins in politics to other fields of social science. Its essence is unexpected speed of change, and this requires a judgment of speed in the context of a longer perspective of historical change, the choice of which is likely to be debatable in the extreme. Leaving the judgmental issue aside for the moment, one could characterize the history of our subject in terms of a series of "revolutions," very broadly defined, as follows. Economics as we know it began with what might be called the "Smithian Revolution" against the established body of doctrines generically described as "mercantilism," a revolution which changed ideas on the nature and sources of the wealth of nations and the policies required to promote the growth of what we now call "affluence." The Ricardian revolution turned the attention of economists from concern with national wealth and its growth to the distribution of income among social classes and the interactions of growth and income distribution. The marginalist revolution of the 1870's essentially introduced a new and superior analytical technology for dealing with Ricardo's distribution problem, in the process gradually depriving Ricardian economics of its social content; hence, the results of that revolution have been described as neo-Ricardian or more commonly neo-classical economics.

Contemporary economics is based on this development and on at least four dis-

cernible "revolutions" that occurred in the late 1920's and in the 1930's. One was the imperfect-monopolistic competition revolution, which challenged the validity of the assumption of perfect competition on which value theory had come to be built following the marginalist revolution, and particularly the conclusions about the welfare effects of competition to which that theory led. This revolution has more or less fizzled out, though its fossilized remnants continue to plague both students and their instructors in elementary courses. Another was the empirical or econometric revolution, with its insistence initially on the measurement of economic relationships and, subsequently and more ambitiously, on the testing of economic hypotheses—though the "testing of hypotheses" is frequently merely a euphemism for obtaining plausible numbers to provide ceremonial adequacy for a theory chosen and defended on *a priori* grounds. The third was the general equilibrium revolution, based on the introduction by Hicks and Allen of the continental Walrasian-Paretoan approach into the Anglo-Saxon tradition in replacement of the then-dominant Marshallian partial-equilibrium approach. Finally, and most sweeping in its effects, there was the Keynesian Revolution in monetary theory.

By contrast with the abundance of revolutions, counter-revolutions are hard to find in the development of economic thought. About the closest one can come to a counter-revolution in the history of economic thought is to interpret the development of the Austrian theory of value as a counter-revolution against the socialist, and especially the Marxist, tradition of economic theorizing; and that aspect of the work of the Austrian school was a side issue in the marginalist revolution. The monetarist counter-revolution of contemporary times is probably the first signifi-

cant counter-revolution in the development of our subject. In venturing this judgment, however, I should note that the disrepute into which the theories of imperfect and monopolistic competition have fallen, as theories of contemporary industrial competition, in the period since the second world war could be described as the result of an intellectual counter-revolution, based on a combination of faith in the preexisting theory of competition and devotion to the empirical revolution; and also that, if one is prepared to disregard the political labels that people choose to attach to themselves, the left-wing student and faculty demand for a politically and socially relevant "radical" economics and protest against emphasis on mathematical and econometric quantification can be classed as counter-revolutionary, inasmuch as it seeks to revert to the pre-marginalist-revolution concern with the economic system as a system of relationships among social classes.

As I have already mentioned, the chief problem in identifying revolutions and counter-revolutions and distinguishing them from slower and more comprehensible and rational processes of change in economic thought is to arrive at a judgment of the relative speed of change and the degree to which the speed is justifiable. From this point of view, some of what I have just now described as revolutions were not really revolutionary—notably the Smithian and marginalist revolutions, the imperfect-monopolistic competition revolution, and the general equilibrium and empirical revolutions. The Smithian and marginalist revolutions spread relatively slowly, through the force of their scientific superiority and intellectual appeal and the process of natural wastage of their opponents. The imperfect-monopolistic competition revolution was the end result of puzzling by many minds over a problem that Marshall had stated but

had been unable to solve satisfactorily—the existence of downward-sloping cost curves for individual firms. The general equilibrium revolution was a result of the delayed appreciation by economists of the need for a better command of mathematical techniques, the delay being occasioned by the long association of the subject with philosophy in the English academic tradition and its continuing association with law in the continental tradition. And the empirical revolution depended on the development of the techniques of statistical inference—most of the historically great economists were quantitatively oriented, or at least paid lip service to the need for quantitative work, but lacked the requisite tools to carry out such work themselves. For real intellectual revolutions, we are left with three major examples: the Ricardian revolution, the reasons for whose rapid propagation were examined some twenty years ago by S. G. Checkland [2], the Keynesian revolution, and the monetarist counter-revolution. These last two are the subject of my lecture today.

My concern, specifically, is with the reasons for the speed of propagation of the monetarist counter-revolution; but I cannot approach this subject without reference to the reasons for the speed of propagation of the Keynesian revolution, since the two are interrelated. Indeed, I find it useful in posing and treating the problem to adopt the “as if” approach of positive economics, as expounded by the chief protagonist of the monetarist counter-revolution, Milton Friedman, and to ask: suppose I wished to start a counter-revolution against the Keynesian revolution in monetary theory, how would I go about it—and specifically, what could I learn about the technique from the revolution itself? To pose the question in this way is, of course, to fly in the face of currently accepted professional ethics, ac-

cording to which purely scientific considerations and not political considerations are presumed to motivate scientific work; but I can claim the protection of the “as if” methodology against any implication of a slur on individual character or a denigration of scientific work.

From this point of view, obviously, the first problem is to identify the elements in the situation at the time of the *General Theory* that accounted for its rapid acceptance and propagation among professional economists. Such elements are of two types, one relating to the objective social situation in which the new theory was produced, the other relating to the scientific characteristics of the new theory itself.

As regards the objective social situation, by far the most helpful circumstance for the rapid propagation of a new and revolutionary theory is the existence of an established orthodoxy which is clearly inconsistent with the most salient facts of reality, and yet is sufficiently confident of its intellectual power to attempt to explain those facts, and in its efforts to do so exposes its incompetence in a ludicrous fashion (on this see [8]). Orthodoxy is, of course, always vulnerable to radical challenge: the essence of an orthodoxy of any kind is to reduce the subtle and sophisticated thoughts of great men to a set of simple principles and straightforward slogans that more mediocre brains can think they understand well enough to live by—but for that very reason orthodoxy is most vulnerable to challenge when its principles and slogans are demonstrably in conflict with the facts of everyday experience.

So it was in the 1930's, and particularly in the 1930's in Britain, which had already experienced a decade of mass unemployment associated with industrial senescence and an overvalued exchange rate, mass unemployment which the prevailing orthodoxy could neither explain nor cope with. This, it may be noted, was

in large part the fault of the economists themselves. There existed already a body of monetary analysis that was quite capable of explaining both Britain's and the industrial world's unemployment problems as a consequence of monetary mismanagement. But, hypnotized by the notion that money is merely a veil cast over real phenomena—the homogeneity postulate of contemporary monetary theory—the economists of the time attempted to explain what were essentially monetary phenomena by real causes. Eminent British economists sought to explain mass unemployment as a consequence of the satiation of real human wants, a satiation that should have produced a general reduction in working hours but unfortunately and inexplicably operated instead differentially to reduce the working hours of a substantial part of the population to absolute zero. Other economists viewed the depression as a punishment justly visited upon enterprises and individuals for past sins of speculation and erroneous microeconomic decision-taking. The concern for microeconomic explanations diverted attention from what the available macroeconomic analysis could have said about the problem; it also led to the recommendation of *ad hoc* remedies such as public works that lacked any firm grounding in theory as generally understood.

In this situation of general confusion and obvious irrelevance of orthodox economics to real problems, the way was open for a new theory that offered a convincing explanation of the nature of the problem and a set of policy prescriptions based on that explanation. Such a theory, however, would have to possess certain characteristics if it were to win intellectual acceptance and political success. In particular, it would have to come from within yet offer liberation from the established orthodoxy—for one must remember that orthodoxy includes both an estab-

lished conservative orthodoxy and an established self-termed "radical" orthodoxy, and, since each recognizes and accommodates the other's arguments, there is no real hope of progress being achieved by a switch from one position to the other.

To be more specific, a revolutionary theory had to depend for its success on five main characteristics—here I must admit that I am conducting my analysis in the blinding light of hindsight. First, it had to attack the central proposition of conservative orthodoxy—the assumed or inferred tendency of the economy to full employment—with a new but academically acceptable analysis that reversed the proposition. This Keynes did with the help of Kahn's concept of the multiplier and his own invention of the propensity to consume. Second, the theory had to appear to be new, yet absorb as much as possible of the valid or at least not readily disputable components of existing orthodox theory. In this process, it helps greatly to give old concepts new and confusing names, and to emphasize as crucial analytical steps that have previously been taken as platitudinous; hence, in the *General Theory*, the marginal productivity of capital became the marginal efficiency of capital; the desired ratio of money to income, the *k* of the Cambridge tradition, became a minor constituent of the new theory of "liquidity preference;" and the *ex post* identity of savings and investment, which previous theorists including Keynes himself had rightly recognized as unhelpful to dynamic analysis, became the *sine qua non* of right reasoning.

Third, the new theory had to have the appropriate degree of difficulty to understand. This is a complex problem in the design of new theories. The new theory had to be so difficult to understand that senior academic colleagues would find it neither easy nor worth while to study, so that they would waste their efforts on pe-

ripheral theoretical issues, and so offer themselves as easy marks for criticism and dismissal by their younger and hungrier colleagues. At the same time, the new theory had to appear both difficult enough to challenge the intellectual interest of younger colleagues and students, but actually easy enough for them to master adequately with a sufficient investment of intellectual endeavour. These objectives Keynes's *General Theory* managed to achieve: it neatly shelved the old and established scholars, like Pigou and Robertson, enabled the more enterprising middle- and lower-middle-aged like Hansen, Hicks, and Joan Robinson to jump on and drive the bandwagon, and permitted a whole generation of students (as Samuelson has recorded) to escape from the slow and soul-destroying process of acquiring wisdom by osmosis from their elders and the literature into an intellectual realm in which youthful iconoclasm could quickly earn its just reward (in its own eyes at least) by the demolition of the intellectual pretensions of its academic seniors and predecessors. Economics, delightfully, could be reconstructed from scratch on the basis of a little Keynesian understanding and a lofty contempt for the existing literature—and so it was.

Fourth, the new theory had to offer to the more gifted and less opportunistic scholars a new methodology more appealing than those currently available. In this respect, Keynes was lucky both in having a receptive audience available, and to hit somewhere conveniently between the old and the newly emerging styles of economic theorizing. The prevailing methodological orthodoxy was that of Marshall—a partial-equilibrium approach set within a clear appreciation of the two complex problems of general equilibrium and of historical change, and hence both unsatisfactory at the simple level of partial-equilibrium analysis taken by itself, and ex-

tremely difficult to apply skillfully in a broader analytical and social context. The new methodological challenge was coming from the explicitly mathematical general-equilibrium approach of Hicks and Allen, an approach whose empirically and historically almost empty generality was of little general appeal. The *General Theory* found a middle ground in an aggregated general-equilibrium system which was not too difficult or complicated to work with—though it demanded a substantial step forward in mathematical competence—and which offered a high degree of apparent empirical relevance to those who took the trouble to understand it.

Finally, the *General Theory* offered an important empirical relationship for the emerging tribe of econometricians to measure—the consumption function, a far more challenging relationship than the demand for sugar, a relationship for which the development of national income statistics provided the raw material needed for estimation, and which could be estimated with surprising success given the limitation of the available data to approximately a single business cycle.

In my judgment, these factors accounted for the success of the Keynesian revolution: on the one hand, the existence of an important social and economic problem with which the prevailing orthodoxy was unable to cope; on the other hand, a variety of characteristics that appealed to the younger generation of that period—notably the claim of the new theory to superior social relevance and intellectual distinction, its incorporation in a novel and confusing fashion of the valid elements of traditional theory, the opportunity it offered to bypass the system of academic seniority by challenging senior colleagues with a new and self-announcedly superior scientific approach, the presentation of a new methodology that made general-equilibrium theory both manageable

and socially relevant, and the advancement of a new empirical relationship challenging for econometricians to estimate.

The very success of the Keynesian revolution, however, ensured that it would in its turn become the established orthodoxy, and as such be as vulnerable as the old to revolutionary attack—which would necessarily have to be a counter-revolutionary attack. Keynes himself, as Leijonhufvud's monumental reinterpretation of his thought [9] has reminded us, had a seasoned and subtle mind, conscious both of the flow of economic history and of the role of theory as an adjunct to policy-making in a given set of historical circumstances. His followers—which means the profession at large—elaborated his history-bound analysis into a timeless and spaceless set of universal principles, sacrificing in the process much of his subtlety, and so established Keynesianism as an orthodoxy ripe for counter-attack.

There are several factors in this transmogrification worthy of note. The first, and probably most important, has been the conviction of Keynesians that the mass unemployment of the 1930's represents the normal state of capitalist society—more accurately, of capitalist society unaided by Keynesian management—and that unemployment is always the most urgent social problem. This view was elevated into a dogma in the United States under the leadership of Alvin Hansen, whose theory of secular stagnation was the subject of his Presidential Address to this Association [6]. While that theory has been quietly forgotten, or frugally converted into a theory applicable to the underdeveloped countries, vestiges of it linger on in the thinking of American Keynesians. The view that unemployment is the overriding social problem also lingers on among British Keynesians such as Joan Robinson, Roy Harrod, and Thomas Balogh, though I should note that Nicholas Kaldor has for many years taken a

much more optimistic view of the resilience of capitalism. The corollary of the Keynesian view of the primacy of the unemployment problem has been a pronounced tendency to play down the adverse economic consequences of inflation, and to assume that, if only the unemployment consequences of anti-inflationary policies were properly understood, society would cheerfully agree to adopt and implement an incomes policy instead.

A second factor in the transformation of Keynesianism into an orthodoxy has been that people who made their academic reputations and earned their present status on the basis of an early and enthusiastic conversion to Keynesianism in the late 1930's and early 1940's have continued to trade on their foresight, to the academic detriment of their juniors, who have never had the same chance to jump onto the front—and not the rear—of an academic bandwagon. This factor has been far more effective in paving the way for a monetarist counter-revolution in the United States, where institutional competition prevents centralized control of professional advancement, than in the United Kingdom, where Oxbridge continues to dominate the academic scene.

A third factor has been that, while the Keynesian revolution in its time offered a tremendous liberation to the energies of young economists in the fields of pure theorizing about concepts, the construction of macroeconomic general-equilibrium models, and the estimation of econometric models of the economy, these activities have run into diminishing returns so rapidly that they have ceased to be appealing to young and ambitious economists.

The result has been that—beginning perhaps sometime in the mid-1950's—Keynesianism has become itself an established orthodoxy, ripe for attack in exactly the same way as what Keynes chose to call “classical economics” and to attack

in the 1930's. It has had the same two vulnerable characteristics: inability to prescribe for what has come to be considered a major social problem—inflation, in contrast to the unemployment of Keynes's time—and a dependence on the authority and prestige of senior scholars which is oppressive to the young. Also, ironically enough in view of Keynes's own long concern with the influence of money on the economy, it has suffered from the same major defect as the orthodoxy Keynes attacked—the attempt to explain essentially monetary phenomena in terms of a mixture of real theory and *ad-hoc*-ery, and specifically to explain inflation in terms of real effective demand and the Phillips curve. The fact that Keynesian economics has stumbled into the same pitfall as the "classical" orthodoxy it succeeded is, perhaps, an indication of the difficulty of monetary theory as contrasted with value theory, as well as of the perils of abandoning monetary theory in favor of what appears seductively to be more reasonable common sense.

If, in accordance with the "as if" methodology of positive economics that I adopted earlier in this lecture, one posed the question of how to mount a counter-revolution against Keynesian orthodoxy, and considered the question in the light of the factors that contributed to the success of the Keynesian revolution, one would, I think, be driven inescapably to two sets of conclusions.

The first would be the need to find an important social problem that the established orthodoxy is incapable of dealing with, even though it tries its best and claims to be successful. The second would be the need to develop a counter-revolutionary theory that had the requisite characteristics to be academically and professionally successful in replacing the previous revolutionary theory.

The obvious answer to the first problem—finding an important social problem

that orthodox theory cannot solve—is to concentrate on the issue of inflation, the issue that Keynesian theory was least well designed to deal with. The trouble with that answer has been that, under the influence of both experienced inflation and Keynesian theory, the public has for the most part not been much concerned about the economic evils of inflation, and so has not regarded inflation as an important test of the intellectual strength of Keynesian orthodoxy. The history of the monetarist counter-revolution has, in fact, been characterized by a series of mostly vain efforts to convince the profession and the public (a) that inflation is an important question and (b) that monetarism can provide an explanation and a policy whereas Keynesianism cannot. Proposition (b) is eminently plausible; but it can only get a hearing if proposition (a) is accepted first; and, aside from a brief interlude in the late 1950's, the public has become convinced of proposition (a) only very recently. It is no accident that the appearance of monetarism as a strong intellectual movement has had to wait until the aftermath of the escalation of the war in Viet Nam in 1965. It is even less of an accident that its current success has depended on a prior Keynesian claim to, and acceptance of, responsibility for efforts to stop inflation by Keynesian fiscal means, under the auspices of the "New Economics." Monetarism has until the past few years been in the position of investing a great deal of intellectual ability in analyzing problems and producing solutions that no one else has considered worth the effort involved. It has eventually become a public force less by its own efforts than as a consequence of the "New Economics" overreaching itself when it was riding high in the formation of national economic policy. The "New Economics" was favored by the opportunity to sell Keynesian policies to meet a Keynesian problem; it encountered disaster

when it tried to sell reverse Keynesian policies to meet a non-Keynesian problem. And the monetarist counter-revolution has been cashing in on that mistake of intellectual strategy.

Nevertheless, on this score of social relevance, the monetarist counter-revolution has had certain factors working in its favor which have enabled it to survive and prosper despite the absence of an overwhelmingly obvious inadequacy of the established Keynesian orthodoxy, for most of the postwar period. One has been that, with the growing professionalization of economics and the expansion of academic support of interest in it, it has become increasingly possible for an issue to be deemed scientifically interesting and worthy of investigation even if the general public displays no visible interest in it. Another has been the rise of the United States to the position of a world power, which has made the exploration of issues of no direct relevance to the economic interests of the United States nevertheless worth pursuing as potentially matters of the national interest in the world economy. Both the hyper-inflations in Europe and elsewhere that followed the two world wars, and the strong inflations that have characterized Latin American economic history, have lent themselves to investigation with the aid of the quantity theory as matters of potential relevance to U.S. economic policy. But, as already mentioned, while these foreign experiences have provided fodder for monetarism, and in the course of time support for the contention that monetarism rests on a far wider base of empirical investigation than Keynesianism, the real counter-revolutionary thrust of monetarism has only developed since inflation became a major problem for the United States itself. Further, it is only since that event—which, given the world importance of the United States, has meant the emergence

of inflation as a worldwide problem—that monetarism has been taken seriously by academic and public opinion in other countries.

Practical social relevance apart, the question of success for a new theory, whether revolutionary or counter-revolutionary, depends on its fitting appropriately into the intellectual climate of its time. Here we may apply what has already been said about the reasons for the successful rapid propagation of the Keynesian revolution to the “as if” question of how to proceed to mount a quantity-theory counter-revolution. There were, I trust you will remember, five elements in the success of the Keynesian revolution, and I shall take them in turn.

The first was a central attack, on theoretically persuasive grounds, on the central proposition of the orthodoxy of the time. In the case of the Keynesian revolution, that proposition was the automatic tendency of the economy to full employment. In the case of the counter-revolution, the obvious point of attack, in a world characterized by high employment and inflationary tendencies, was the vulgar Keynesian orthodox position that “money does not matter.” As James Tobin has pointed out, there is a world of difference between two alternatives to this proposition, namely, one, “money does too matter,” and, two, “money is all that matters.” But this difference was easily and conveniently blurred, to the benefit of the counter-revolution, by seizing on the extreme Keynesian position that money does not matter at all as the essence of the prevailing orthodoxy.

The second aspect of Keynesian success was the production of an apparently new theory that nevertheless absorbed all that was valid in the existing theory while so far as possible giving these valid concepts confusing new names. This was the technique followed—again I would emphasize

the "as if" character of my interpretation—in Friedman's classic restatement of the quantity theory of money [3]. The restated quantity theory is, as Patinkin has recently pointed out, essentially a generalization of Keynes's theory of liquidity preference on the basis of a more sophisticated analysis of the nature of wealth and the relation of wealth to income. Novelty and the requisite intellectual confusion were provided by the substitution of the concept of "permanent income" for that of wealth, and the dragging across the trail of the red herring of human capital that was emerging from other work being conducted at Chicago at that time. Nevertheless, the restatement of the quantity theory of money did include one important and genuinely novel element, drawn not from Keynes but from his predecessors in monetary theory, which was highly relevant to the problem of inflation and which continues to distinguish quantity theorists from Keynesians; this consisted in its emphasis on the Fisherian distinction between the real and the money rate of interest and on the expected rate of price inflation or deflation as determining the difference between the two.

For the reasons just given, the restatement of the quantity theory provided a new theory meeting the third criterion for success, a degree of difficulty of understanding just sufficient to deter the old and to challenge and reward the young, and hence to reopen the avenues of professional opportunity for the ambitious.

The fourth criterion for success was a new and appealing methodology. Here the counter-revolutionary theory could appeal against the tendency of Keynesian economics to proliferate into larger and yet larger models of the economic system, a tendency which sacrificed theoretical insights to the cause of descriptive realism and which had the incidental but important detractors of demanding large sums

of scarce research money available only to senior economists and of turning young economists into intellectual mechanics whose function was to tighten one bolt only on a vast statistical assembly line, the end product of which would contain nothing that could be visibly identified as their own work. In place of this approach, the counter-revolution set up the methodology of positive economics, the essence of which is not to pursue descriptive realism as represented by the largest possible system of general equilibrium equations, but to select the crucial relationships that permit one to predict something large from something small, regardless of the intervening chain of causation. This methodology obviously offered liberation to the small-scale intellectual, since it freed his mind from dependence on the large-scale research team and the large and expensive computer program.

The fifth criterion for success was the advancement of a new and important empirical relationship, suitable for determined estimation by the budding econometrician. That relationship was found in the demand function for money, the stability of which was claimed to be the essence of the traditional quantity theory of money. Presentation of the stable demand function for money as the essence of the quantity theory offered a close parallel to the Keynesian consumption function of the 1930's—a statistical relationship simple to understand theoretically and not too hard to estimate statistically, which promised, nonetheless, to contribute importantly to the resolution of central theoretical issues. Moreover, since intelligent and gifted young men and women will persevere until they succeed in finding statistical validation of an allegedly important theoretical relationship, and will then interpret their results as evidence in favor of the theory that originally suggested the relationship, their efforts will

inevitably be extremely favorable to the theory in question. And so it has proved. A stable demand function for money is by no means inconsistent with the Keynesian macroeconomic general equilibrium model, and indeed is presumed to exist in the construction of the standard IS-LM diagram. But the empirical finding of the existence of such a function has been widely adduced in support of the quantity theory as against the rival Keynesian theory, a procedure justified only by the identification of the Keynesian orthodoxy with the proposition that money does not matter and that velocity is either highly unstable or infinitely interest-elastic.

The quantity-theory counter-revolution could therefore make use of the same factors as facilitated the rapid propagation of Keynesian economics—the attack on a central and widely held theoretical proposition, the development of a new theory that absorbed and rechristened the best of the old, the formulation of that theory in terms that challenged the young and enabled them to leapfrog over the old, the presentation of a new methodology that made more immediate sense than the prevailing methodology, especially in terms of accessibility to the young and to those outside the established centers of academic excellence, and a new and presumptively crucial empirical relationship suitable for relatively small-scale econometric testing.

A counter-revolution, however, has to cope somehow with a problem that a revolution by definition can ignore—though it can trade on it in its propaganda—the problem of establishing some sort of continuity with the orthodoxy of the past. Specifically, the monetarist counter-revolutionaries were burdened with the task of somehow escaping from the valid criticisms of the traditional quantity theory, which the Keynesian revolution had elevated into articles of dogma and self-justi-

fication. These criticisms were, first, that the quantity theory had assumed an automatic tendency to full employment, which was manifestly in conflict with the facts of experience; and, second, that velocity was a highly unstable variable, useful, if at all, only for the *ex post* description of historical events. The restatement of the quantity theory met these criticisms by two counter-contentions: that the question of whether the economy responds to monetary impulses by price-level or by output changes is an empirical question falling outside the domain of monetary theory properly defined, because the quantity theory is a theory of the demand for money and not a theory of aggregate response to monetary change; and that the essence of the quantity theory as a theory of the demand for money is not presumptive constancy of velocity but the stable functional dependence of velocity on a few major variables. The former counter-contention freed the quantity theory from the charge that it was too silly to be worth considering, and opened the way for fruitful scientific controversy and development in monetary theory—though, as I shall explain later, the abnegation of responsibility for explaining the division of the effects of monetary change between price and quantity movements has subsequently proved a serious short-coming of the counter-revolution, now that the counter-revolution has come to be taken seriously. The latter counter-contention, involving emphasis on the existence of a stable demand function for money, permitted the absorption of the best of Keynesian ideas into the quantity theory cause, without any recognized need for acknowledgment of their source. The problem in the case of both counter-contentions was to establish a plausible linkage with pre-Keynesian orthodoxy.

The solution to this problem was found along two lines. The first was the inven-

tion of a University of Chicago oral tradition that was alleged to have preserved understanding of the fundamental truth among a small band of the initiated through the dark years of the Keynesian despotism. The second was a careful combing of the *obiter dicta* of the great neo-classical quantity theorists for any bits of evidence that showed recognition (or could be interpreted to show recognition) of the fact that the decision to hold money involves a choice between holding money and holding wealth in other forms, and is conditioned by the rates of return available on other assets.

Don Patinkin has very recently—and over-belatedly, from the standpoint of the history of economic thought—exploded these efforts to provide bridges between the pre-Keynesian orthodoxy and the monetarist counter-revolution [10]. He demonstrates conclusively that in their theorizing the neo-classical theorists did assume a tendency to automatic full employment, and that in their analyses of practical policy problems they regarded the inherent instability of velocity as a major disturbing element and made no use whatever of the functional relationship between velocity and other aggregate variables implied by their own *obiter dicta*. And he shows specifically that the Chicago quantity theorists—Simons and Mints—were no different from their quantity theory colleagues elsewhere in these respects. There was no lonely light constantly burning in a secret shrine on the Midway, encouraging the faithful to assemble in waiting for the day when the truth could safely be revealed to the masses; that candle was made, and not merely lit, only when its light had a chance of penetrating far and wide and attracting new converts to the old-time religion.

Nevertheless, one should not be too fastidious in condemnation of the techniques of scholarly chicanery used to promote a

revolution or a counter-revolution in economic theory. The Keynesian revolution derived a large part of its intellectual appeal from the deliberate caricaturing and denigration of honest and humble scholars, whose only real crime was that they happened to exist and stand in the way of the success of the revolution. The counter-revolution had to endow these scholars, or at least their intellectual successors, with a wisdom vastly superior to what their opponents had credited them with. *Obiter dicta* and an oral tradition are at least semilegitimate scholarly means to this polemical end. Moreover, as time has passed and the counter-revolution has acquired increasing academic respectability, it has become increasingly possible to admit, and even to brag, that the useful ideas have been drawn from the revolution and not from the preexisting orthodoxy. Indeed, this is a necessary element in a successful counter-revolution, an element for which a previously successful revolution inevitably provides the foundations—because it ultimately becomes possible to draw an intellectually acceptable distinction between the sophisticated ideas of the revolutionary leader and the unsophisticated ideas of the revolutionary followers and executors, and to absorb the former into the counter-revolutionary ideology while discarding the latter as beneath intellectual contempt. The service of drawing this distinction in intellectually acceptable terms has been performed for the monetarist counter-revolution with great scholarly distinction by Axel Leijonhufvud's book on Keynesian economics and the economics of Keynes.

I have in this lecture been concerned primarily with the intellectual and social factors that make it possible to launch a successful revolution or counter-revolution in economic theory. However, I would judge that the key determinant of success or failure lies, not in the academic

sphere, but in the realm of policy. New ideas win a public and a professional hearing, not on their scientific merits, but on whether or not they promise a solution to important problems that the established orthodoxy has proved itself incapable of solving. Keynes, and many other economists in Britain and elsewhere, spent much time in the 1920's and 1930's advocating public works as a cure for unemployment—a cure that, because it conflicted with prevailing orthodoxy, was unacceptable. The *General Theory* was successful, precisely because, by providing an alternative theory to the prevailing orthodoxy, it rationalized a sensible policy that had hitherto been resisted on purely dogmatic grounds. Similarly, the monetarist counter-revolution has ultimately been successful because it has encountered a policy problem—inflation—for which the prevailing Keynesian orthodoxy has been able to prescribe only policies of proven or presumptive incompetence, in the form of incomes or guidelines policy, but for which the monetarist counter-revolution has both a theory and a policy solution.

No particular point would be served in a lecture of this kind by recounting the stages of accomplishment in the monetarist counter-revolution (see [7]). The advance from strength to strength is summarizable in a few key phrases: the restatement of the quantity theory, a statistical illusion in the judging of Keynesian models, velocity versus the multiplier in U.S. monetary history, monetarism versus fiscalism, and “the new new economics.” The question of interest is whether the monetarist counter-revolution will sweep the board and become the orthodoxy of the future, itself ripe for attack by a new revolution, or whether it will gradually peter out.

Personally, I expect it to peter out, for two reasons. The first, and most impor-

tant, is that I believe the Keynesians are right in their view that inflation is a far less serious social problem than mass unemployment. Either we will vanquish inflation at relatively little cost, or we will get used to it. The odds at present are that we will accept it as a necessary price of solving other pressing domestic issues—this seems to be the current view of the present Administration—and in that case monetarism will again be reduced to attempting to convince the public of the importance of the problem it is equipped to solve before it can start arguing about the scientific superiority of its proposed solution to the problem. The second reason is that monetarism is seriously inadequate as an approach to monetary theory, judged by prevailing standards of academic economics, and in the course of repairing its intellectual fences and achieving full scientific respectability it will have to compromise irretrievably with its Keynesian opposition.

The most serious defects of the monetarist counter-revolution from the academic point of view are, on the one hand, the abnegation of the restated quantity theory of money from the responsibility of providing a theory of the determination of prices and of output, and, on the other hand, its continuing reliance on the methodology of positive economics. Abnegation of responsibility for analyzing the supply response of the economy to monetary impulses, and particularly the disclaiming of the need for an analysis of whether monetary changes affected prices or quantities, was, as I have explained earlier, necessary to the restoration of the quantity theory to a position of academic respectability. But this need was transitory: once the quantity theory regained academic respectability, it was obliged to resume responsibility for the short-run forecasting of aggregate movements of prices and

quantities (see [5]). This it has begun to do, most importantly through the research work of the Federal Reserve Bank of St. Louis, and with appreciable success; but it has been lured into playing in a new ballpark, and playing according to a different set of rules than it initially established for itself.

In similar fashion, the methodology of positive economics was an ideal methodology for justifying work that produced apparently surprising results without feeling obliged to explain just why they occurred, and in so doing mystifying and exciting the interest of noncommitted economists and wavering Keynesians. But the general equilibrium and empirical revolutions of the recent past have taught economists to ask for explicit specification of the full general equilibrium system with which the theorist or empiricist is working, and to distrust results that appear like rabbits out of a conjurer's hat—and an old-fashioned top hat at that. The demand for clarification of the mechanism by which results can be explained is contrary to the methodology of positive economics, with its reliance on the “as if” approach. But it will have to be answered satisfactorily if the monetarist counter-revolution is to win general acceptance among the profession; and the attempt to answer it will necessarily involve the counter-revolutionaries in the opposing methodology of general-equilibrium systems and multi-equation econometric models. The quantity theorists have already begun to extend their efforts into simultaneous-equation formulations and estimations of economic relationships. In so doing, they have been making important methodological compromises with the Keynesian opposition—or, to put it another way, reaching out for a synthesis between the revolution and the counter-revolution.

In summary, it seems to me that the

monetarist counter-revolution has served a useful scientific purpose, in challenging and disposing of a great deal of the intellectual nonsense that accumulates after a successful ideological revolution. But its own success is likely to be transitory, precisely because it has relied on the same mechanisms of intellectual conquest as the revolution itself, but has been forced by the nature of the case to choose a less important political issue—inflation—to stand on than the unemployment that provided the Keynesian revolution with its political talking point, and has also espoused a methodology that has put it in conflict with long-run trends in the development of the subject. If we are lucky, we shall be forced as a result of the counter-revolution to be both more conscious of monetary influences on the economy and more careful in our assessment of their importance. If we are unlucky (those of us who are not good at jumping on band-wagons) we shall have to go through a post-counter-revolution revolution as the price of further progress on the monetary side of our science.

REFERENCES

1. Brunner, Karl, “The ‘Monetarist Revolution’ in Monetary Theory,” (mimeographed, 1970).
2. Checkland, S. G., “The Propagation of Ricardian Economics in England,” *Economica*, New Series, Vol. 16, No. 61 (February 1949), pp. 40–52.
3. Friedman, Milton, “The Quantity Theory of Money—A Restatement,” in M. Friedman (ed.), *Studies in the Quantity Theory of Money* Chicago 1956, pp. 3–21.
4. ———, *The Counter-Revolution in Monetary Theory*, London 1970.
5. ———, “A Theoretical Framework for Monetary Analysis,” *Jour. P. Econ.*, 78, No. 2 (March-April 1970), pp. 193–238.
6. Hansen, Alvin, “Economic Progress and Declining Population Growth,” *A.E.R.*, March 1939, 29, pp. 1–15.

CURRENT STATUS OF INCOME MAINTENANCE EXPERIMENTS

The Graduated Work Incentive Experiments: Current Progress

By HAROLD W. WATTS
University of Wisconsin

The experimental projects for estimating labor supply and other responses to negative-tax-type transfers have been reported at the last two annual meetings. Since last year the projects sponsored by OEO have been in full operation and are maturing toward the end of their scheduled run. New projects have been authorized by HEW with the objective of complementing and, in part, replicating the OEO studies. These will be reported separately in this session and I will concentrate on a review of the current status of the experiments that are supervised by the Institute for Research on Poverty. These are (1) the urban experiment (managed by MATHEMATICA) in New Jersey and Pennsylvania and (2) the rural experiment operated directly by the Institute under the direction of Professor Lee Bawden.

The first experiment site was Trenton, New Jersey, where families were enrolled in August 1968. Only eight months of the scheduled three-year experimental period for Trenton remain, and we are in the process of completing plans for ending the payment and interviewing programs for the small "pilot" sample in that city. Our objectives in this phase-out are both to soften the impact of payment withdrawal and to obtain useful data from the tail end of the experimental period and from the

post-termination period. We regard it as very important that the experiment be ended without a "bad taste" in the mouths of experimental subjects, the scholarly community, or the operational staff, in order to preserve the possibility of other social experiments in the future.

We expect to learn useful lessons in Trenton for the phase-out of the other cities. Paterson-Passaic will complete three years in February of 1972, Jersey City in June and Scranton in September of that same year. Consideration is being given to extending a portion of the sample for a four-year period (or perhaps longer) to provide a basis for evaluating the consequences of the temporary nature of the experiment as contrasted with the presumptive permanence of a "real" program. If this is done, it will certainly take place in the sites that still have the longest period to run.

The rural experiment is operating in two areas: Duplin County, North Carolina, and Calhoun and Pocahontas counties in Iowa. The Iowa sample completed its first year in November and the group in North Carolina is one month behind. The projected three-year run of this experiment will therefore take it to the end of 1972 plus the phase-out period.

Both experiments are well established at this time and are operating on a more

or less routine basis. Income reports come in every four weeks, benefits are paid every two weeks, interviews are conducted every three months, and benefit levels are adjusted for cost of living every year. Coding, punching, editing and cleaning up the large bodies of accumulating data have proven to be a very challenging problem.

Approximately half the data in the urban experiment has been collected at this time, and about a third of the data in the rural operation. The time is approaching, then, when major analytic studies using these data can be carried out. Indeed, when one considers that the middle year will yield the least biased data (least contaminated by start-up problems and anticipation-of-termination problems), we can say that our "best" data will all be in very soon. These studies, dealing with the main labor supply objective of the experiment, will be carried out during the coming year. It is expected that at this time next year we will have important analytic contributions to make.

In order to provide for exploitation of the data base both at Wisconsin and at MATHEMATICA, concurrently with the data collection process, we have adopted a somewhat unusual data system. After each incremental batch of data has been edited and cleaned up to the point of usefulness for analytic purposes, that file is added to the data available at a central processing point at Stamford, Connecticut. These data may then be manipulated on command from remote teletype terminals in Madison and Princeton. We feel we shall thus avoid both the delays involved in keeping track of duplicate copies of tapes and also, hopefully, the errors and confusions which may come from further modification of the data as problems are uncovered during their use. Less complex but still quite elaborate and essentially

new systems are being put into use to manage the data produced in the rural experiment.

Although there are no major studies to be reported at this time there are a few minor analyses that are of some interest. The first of these is the very preliminary report which appeared last March, based on the first year's operation for half of the total sample. In response to a very urgent plea from the Administration for evidence that would support its legislative program of welfare reform, the recently rejected Family Assistance Plan, some very preliminary results were untimely ripped from the first available batches of data. A somewhat frantic scramble to handcode and punch selected data from the complete questionnaire files was made and these were rapidly tabulated for simple and crude analysis. It is both fortunate and somewhat gratifying that, in spite of the initial haste of this work, subsequent careful checking uncovered no errors which changed the interpretation of the findings. The findings, as many of you may know, were dominated by repeated acceptance of the null hypotheses that no effect on the earnings of either the family head or his spouse is produced by the experimental treatments—i.e., the receipt of income-conditioned transfers using a variety of marginal tax rates and basic support levels. Notwithstanding the weakness of the analysis (the relatively scanty data base was further weakened in its power by the use of quite crude analytic tools), the findings are still of potential value. We can certainly conclude, at the very least, that transfer payments of the kind administered in New Jersey do not induce immediate and widespread withdrawals from the labor force.

It might be said, of course with hindsight, that no one should have expected such an outcome in the first place. Never-

theless, there has been and remains a very wide range of both public and professional opinion about the labor force response to such payments. Even these very preliminary results have in a small way begun to provide evidence that can narrow that diversity of opinion. On the basis of these findings, we can assert, I believe, that the response is not at all in the nature of a mass exodus from the ranks of the gainfully employed, and that if there is to be an obvious and substantial reduction in labor supply, it will have to show up with a considerable lag. It is possible certainly to appeal to the many features of the experiment that are special—it is limited to specific urban northeastern areas, the time period is relatively short, etc. We are certainly administering a program which is different in many respects, no doubt, from a program which might be instituted eventually by the government. But there is a substantial contribution to the fund of relevant evidence in these early results.

The early results, indeed, have received a substantial amount of attention. They were displayed prominently as part of the early presentation of the Family Assistance Plan to the Congress. It is, of course, very difficult to assess how much impact they have had. Probably the best evidence lies in the fact that the study was regarded as important enough to warrant attack by opponents of the President's program. It is at least questionable whether preliminary results at such an early stage of the experiment should have been drawn off and used as a part of a major policy debate. With hindsight again, it is clear that life would have been easier for the experimental staff if the request for early results could have been resisted. At the same time, one cannot expect the process of public policy formation to wait upon the completion of scientific studies. Where ongoing studies can make a sub-

stantial and important contribution to the basis for public decisions, it would seem irresponsible (even if it were possible) to withhold the evidence.

One of the questions raised in the process of challenging the reliability of our findings was the problem of sample attrition. This is, of course, a problem generally encountered in any kind of survey work which attempts repeated interview measurements or other contact with an initially selected sample of families. Panel studies involving more than a dozen contacts over a three-year period have generally experienced quite substantial losses from their original sample. People move from their original dwellings and, either by neglect or design, fail to leave a trail that can be followed for subsequent interviews. Others simply get fed up with the process and refuse further cooperation.

Such panel mortality was anticipated for this study, and expectations about how many families would be lost in the control and experimental groups were built into the original design. In other words, we aimed at having roughly optimal numbers of families in various experimental and control cells at the end of the experiment after allowing for anticipated attrition rates. Last summer a review of our attrition experience was undertaken. The results are displayed below.

Table 1 summarizes the most recent available information on attrition and utilization of welfare for our sample families. We have shown separately our figures for the first half of the sample (families in Trenton, Paterson and Passaic) and for the second half (Jersey City and Scranton). Control and experimental groups are also divided, and the latter are subdivided again into three groups—a low-benefit group, with maximum benefits at 50 percent of poverty; a middle group, with maximum benefits at 75–100 percent of

TABLE 1—CURRENT STATUS OF FAMILIES

	Original Families (no.)	Families Lost (%)	Families Added by Splits (no.)	Families Now Filing (no.)	Families on Welfare (%)
A) Trenton, Paterson and Passaic:					
Low Benefit	83	16.9	5	73	34.2
Medium Benefit	250	17.6	6	211	27.5
High Benefit	32	3.1	2	32	6.2
Experimental	365	16.1	13	316	26.9
Control	285	16.3	26	265	18.5
Grand Total	650	16.2	39	581	22.9
B) Jersey City and Scranton:					
Low Benefit	46	4.3	1	45	13.3
Medium Benefit	211	2.4	1	208	12.5
High Benefit	103	4.8	0	98	9.2
Experimental	360	3.3	2	351	11.7
Control	346	9.2	6	329	14.3
Grand Total	706	5.0	8	680	12.9
C) Entire Sample:	1356	10.3	47	1261	17.5

poverty; and a high-benefit group, with a maximum benefit of 125 percent. Losses to date are predictably higher in Trenton-Paterson-Passaic because they have been operating longer and also because our anti-attrition measures were adopted after these cities were already in operation. Also, the majority of our Spanish-speaking population is in Paterson-Passaic. The phenomenon of frequent moves back to Puerto Rico explains the high attrition. Our guesses about final attrition suggest that it will be about twice as high in Trenton-Paterson-Passaic as in Jersey City-Scranton.

The third column of the table shows additions to the number of currently active families that have been produced by subdivision of originally enrolled families. These account for the difference between the original sample as reduced by attrition and the size of the current sample. (Re-

maining discrepancies come from families who were lost after they had split.)

We found then that we had lost just over 10 percent of the families; but at that point, because of split families, the number of currently active families was equal to 93 percent of the number that started the program. By extrapolating the experience to date for various groups in each of the several sites up to the end of the three-year period, we estimate that the eventual losses will be between 20 and 25 percent of the original families. This could, again because of offsetting additions coming from family splits, leave us with 85 percent or so of the original sample size.

As mentioned above some anti-attrition measures were taken to prevent the attrition from being somewhat higher. One measure which seems to have been successful was to increase money payments

to the control families to motivate their continued cooperation. Besides paying for their time when they are interviewed each quarter, we make an \$8 monthly payment in return for a current address slip. This enables us to keep closer track of families that move and also provides a more continuous kind of financial reinforcement for their continuation in the panel. Similarly, the payment made to experimental families whose income has risen above the break-even point, which qualifies them for transfer payments, has been increased. These families are expected to file income report forms every four weeks even if they are not qualified for the transfers. They now receive a \$20 payment every four weeks to prevent the loss of what are in many ways the most interesting observations—i.e., those whose income goes above or merely stays above the break-even level.

These measures explain, at least in part, the better attrition experience in the cities most recently added to the experiment. In the final analysis, of course, the damage to the evidence caused by attrition will be estimated. This certainly depends upon the nature of the observations that were lost. We expected, and experienced, greater losses both for control families and for families that receive very small or zero benefits from the experimental treatment. Families that receive quite large benefits understandably stick with us. But we need to probe much more deeply into the question of whether the losses bias the sample heavily in one direction or the other because of differential losses by age, education, ethnic group, or other crucial categories.

It must be remembered that, at least for the attrition noted in Table 1, we do have substantial amounts of information about the families that were lost. They all completed a 90-minute enrollment interview in order to be included in the initial sam-

ple. And, again, not all of those families were lost at the beginning of the experiment, so we shall have partial data that can be used in the assessment of potential bias. Families that are lost near the end of the experiment are not losses at all, except for analysis that requires data extending to the very end of the experiment.

Another kind of loss is also shown in Table 1. These are families now enrolled in public welfare programs—for the most part in the Aid to Families with Dependent Children category. This complication was only partly anticipated when the experiment was planned. Because New Jersey did not have the so-called unemployed parent provision in the Aid to Dependent Children program, it was hoped that an experiment dealing with families that included a potential male wage earner (usually the family head) could be carried out with a minimum of "contamination" from sample families' participation in public welfare programs. This situation lasted only until January of 1969, the beginning date for the Unemployed Parent provision of AFDC in New Jersey.

Many of our low income control families, as well as our experimental families, became eligible for benefits under that program. Moreover, in many cases families eligible for AFDC benefits could receive higher payments than they were entitled to under the plan assigned to them in our experiment. We have adopted the rule that a family may freely choose the public assistance payments or our experimental payments, depending on its preference or advantage, but that it may not receive both simultaneously. Moreover, such families can reverse their choice at any time, or as many times as they wish. We do, however, make every effort to obtain the same interview and income report information from the families that choose to go on welfare.

Analytically, the introduction of wel-

fare eligibility for some part of our sample introduced a complication, but by no means a fatal one. The availability of welfare is uniform for control and experimental families alike. Obviously, if the welfare benefits uniformly dominated in attractiveness the experimental treatments, all families would choose them and there would be no discernible effect of the treatments themselves. But insofar as the public assistance program can be described in the same way as the experimental treatments—i.e., in terms of a basic support level of guarantee and a marginal tax rate—it is also possible to regard welfare as an additional optional treatment. This again would permit relatively unadulterated estimates of the effects of our treatments, as well as the welfare treatment, so long as the welfare treatment is not chosen by the overwhelming majority of the families. So far there is little in our experience to suggest that this is a serious threat.

It is not improbable that by the end of the project as many as 30 or 40 percent of the control families will be enrolled in AFDC and some smaller fraction of the experimental families. Part of the families on welfare, however, will come from changes in family status after the experiment started. In other words, a family that started out with a male wage earner may lose him in mid-experiment and thus make the family eligible for the Aid to Dependent Children program—and would have done so even if the law had not been changed in New Jersey. It should be noted parenthetically that our accounts of families on welfare are much more accurate for experimental families with whom we have at least monthly contact than it is for control families. The latter are contacted only every three months and, since they are receiving no experimental benefits, it is not administratively crucial to determine whether or not they are on welfare.

Attrition in the first six months of activity in the rural experiment has been quite light—only 2 percent had been lost by June, 1970. It is far too early to conjecture about the continuation of that very low rate and there are just too few cases to attempt any verification of the tendencies noted in New Jersey. It can, however, be speculated that residents of rural areas in general have less anonymity and it is just less likely that they will disappear without a followable clue.

Level of Understanding

One subject of potential importance has been explored in the rural experiment. This is the question of the level of understanding of the experimental treatment on the part of the families. This was undertaken as a part of a re-enrollment process carried out eight months after the program was initiated. Re-enrollment took the form of an interview which examined the families' understanding of the particular treatment (i.e., negative tax plan) that had been assigned to them. The examination was informally scored by the interviewer who then proceeded to correct the family's understanding and to clarify areas that were not covered at the original enrollment. In general, the experience at enrollment time suggests that many families do not listen very closely after they have been presented with their first benefit check. In addition to the verbal description presented at that time, they are given an abbreviated copy of the rules. (A complete copy is available in the field office on request.) They are also encouraged to seek clarification whenever they are uncertain either about their reporting requirements or about the size of the payments they have received. Nevertheless, after eight months of operation, there were quite a few families that were mistaken about seemingly critical aspects of the program. Approximately three-fourths understood that they could continue in the

program if they lost (or quit) their job, or if they moved out of the county, leaving 25 percent wrong about this. Only about half of the families gave the appropriate response ("don't care") when asked what the program operations wanted them to spend their money on. On the other hand, 85-90 percent of the families knew that the experiment would last only for three years.

While a narrow majority of the Iowa families knew roughly the size of the income guarantee or maximum benefit, substantially less than half the families in North Carolina had anything like an accurate picture of their potential benefits if their income fell to zero. Very comparable results were obtained from questions relating to the breakeven point. Apparently as many as 30 percent were not clear that their benefits would fall when their income rose and a very few families responded correctly to a somewhat tricky question about the inverse relationship between benefit size and total disposable income.

In many ways these results are very disappointing—and indeed discouraging—if one assumes that an appropriately representative response depends upon a complete understanding of the treatments. But, in fact, norms of understanding of other operating programs are not readily available. We don't know, for example, how well the Internal Revenue Code is understood by the average taxpayer. Presumably an experiment of this kind should aim at a level of knowledge on the part of participants that is equivalent to the understanding that would be attained in a real program. Unfortunately, no one knows precisely what or how much knowledge this involves. There is (in a more detailed breakdown of the rural survey) some evidence that those persons who might be presumed to need to understand a particular feature, because of its pertinence to decisions they regularly make,

are more apt to have such knowledge in their head. For instance, more of those families with a third adult know what special provisions relate to such people than the rest of the families do. Given that the families did receive further instruction about the nature of the program, it will be of some interest to see if a subsequent test will reveal any increase in understanding.

Conclusion

During the coming year an increasing part of this sample will have recorded the response for the crucial middle, or second year, of the experiment. Data from this period should be relatively free of any distortions coming from lag adjustments at the beginning or from anticipation of payment withdrawal at the end. It is at this point that heavy economic analysis and econometric work can begin. First priority will be given to estimation of the crucial parameters of the labor supply function. With these it should be possible to estimate the effect on aggregate labor supply (or at least some major parts of it) of instituting some kind of cash transfer system affecting the working poor. In a data source as rich as this one will be, it is possible to take many different approaches and use many different models for analyzing the data. We intend, indeed, to include in the principal report from the study the results from several different approaches, along with some reconciliation of any differences that appear. But it is also hoped that subsequently others will be attracted to using these same data for further exploration of labor supply and other questions. We expect that by this time next year it will be possible to present the first of the major findings from the study. I am sure this will be gratifying to the many people who have shown a continuing interest in this undertaking, and who have up to now had to be content with the sequence of progress reports of which this paper is a part.

The Seattle Experiment: The Combined Effect of Income Maintenance and Manpower Investments

By MORDECAI KURZ, *Stanford University*

and

ROBERT G. SPIEGELMAN, *Stanford Research Institute*

The income maintenance experiments now in process comprise the most ambitious attempt at social experimentation ever undertaken in this country. The first income maintenance experiment was launched by the Office of Economic Opportunity in New Jersey about two years ago. In this experiment approximately 600 two-parent families were provided financial support. The New Jersey design has been carried into two rural areas in a second OEO experiment. An experiment undertaken by HEW in Gary, Indiana, will extend the population coverage to one-parent families, but will deal only with black families. It will add a day-care treatment to the experimental design.

The Seattle experiment extends treatment to all family units with heads between the ages of 18 and 58, black and white, including families with one and two heads. Spatially, it adds the western United States to the area covered by experiments. Conceptually, it adds an experimental manpower program to the design. The manpower treatment is intended to counteract the expected negative impact of income maintenance on work effort.

The purpose of this paper is to present the basic design of the Seattle experiment. Since the experiment is in the initial stage of implementation, we shall restrict ourselves to a general description of the design, and the problems that we have en-

countered and expect to encounter in the future.

The Income Maintenance Component

To launch a viable experiment, it is necessary to offer support levels that are at least as high as the alternative effective support levels in existence. For able-bodied adults under 65, the most important alternative support program is Aid to Families with Dependent Children (AFDC). This program offers support primarily to unemployed one-parent families, although two-parent families are eligible if the father is unemployed and actively seeking work. A small number of families with a partly employed parent also receives support. The effective level of welfare support for a typical family of four in Seattle with no other means is approximately \$3,400, although some families with older children receive as much as \$4,000. Thus, a \$3,600 support level replaces welfare for the unemployed family with a program that offers less categorical and demeaning aid rather than aid at a significantly higher level of financial support. However, the working poor, for whom AFDC is essentially unavailable, will materially benefit from an income maintenance program even at the \$3,600 level. In the experiment, there will be a second support level at \$4,800 a year, introduced to enable the impact of higher

support on the various parameters to be measured. It should be noted that the \$3,600 and \$4,800 support levels refer to that available for a family of four. A family-sized index is used to normalize the level for smaller and larger families.

Turning now to the problem of the tax structure, we note first that in New Jersey only constant tax rates of 30%, 50%, and 70% have been used. In the final evaluation, it is intended to estimate the marginal response function using these three observed tax rates. It is our opinion that using such linear approximations has significant deficiencies. As a result, the Seattle experiment uses a two-parameter tax function, in which Y is taxable income, and the average tax rate is

$$T(Y) = T_1 - RY.$$

This permits use of linearly declining tax rates, where T_1 is "the initial tax rate," and R is "the rate of decline of the average tax rate." In the experiment, T_1 will take the values 50%, 70%, and 80%, and R will take values of 0 and 0.025 per \$1,000 of taxable income. Thus, constant tax rates will remain an essential ingredient of the Seattle experiment, but a new dimension will be added in those cases in which R is not equal to zero. The reasons for considering declining tax rates are as follows:

- (1) When a real program is introduced, the negative taxes in the income maintenance program and positive tax rates in the regular federal income tax programs will probably be combined. If the rate of positive tax is zero or 14% around the cutoff¹ (or ceiling), then an integration of the negative and positive tax structures will call for the negative tax to be declining

toward zero from the high initial rates. Thus, the declining tax rate is a more realistic structure.

- (2) From the point of view of incentives, a declining tax rate has the advantage that, once the individual decides to work at all, his incentive to increase his work effort is enhanced by a declining marginal tax rate. That is, if the high initial tax does not discourage a person from entering the work force, then the declining tax system tends to push him toward full employment.
- (3) From the social cost viewpoint, a declining tax structure with a high level of initial tax is a cheaper way of achieving the social goals of the programs. It provides the basic support to those who are in greatest need, but gets them off the program very quickly with a high level of incentives to increase their work effort once the decision to work is made.
- (4) The declining tax structure and its consequences as described above are consistent with the experimental manpower program. We view the program essentially as one that is designed to achieve two purposes: (a) to provide individuals with the incentive "to go over the hump" and join the labor force, and (b) once in the labor force, through the combination of declining tax rates and increased training, to provide a double incentive to raise their labor participation.
- (5) From the experimental viewpoint, it is not clear that individuals think only in terms of their marginal tax rates and remuneration. Households may consider radical changes in their behavior due to the introduction of income main-

¹ The cutoff is the point at which income maintenance payments are equal to zero.

tenance and the linear approximation proposed by constant tax rates may not provide adequate information.

The declining tax structure will provide observations on many marginal tax rates and may test the response of individuals to different regimes. A family that may consider large changes rather than marginal changes may respond to the regime of declining tax rates differently than to a regime of constant tax rates. Response to "a regime" per se may be viewed as irrational behavior; however, in relation to problems of taxation, there is no way of ruling out such effects, and for this reason we expect to consider this issue with great care.

The introduction of declining tax rates does introduce econometric and substantive complications. For example, when considering families of different sizes, we have introduced a family-size index to adjust the support level. Thus, in order to compare the impact of marginal tax rates, the tax function cannot be applied to the nominal income of the family, but rather it must be applied to an income which provides equivalent per capita purchasing power to that provided a family of four members.

The Manpower Component of SIME

The principal objective of the experimental manpower program is to determine, first, the price elasticity of demand for vocational training, and second, the effect of investment in human resources on work effort.

The recent experience with manpower programs for the poor in the United States is far from encouraging, and this feeling has led us back to essentially a textbook

concept of experimentation. In essence, the various manpower treatments will consist of training subsidies in the form of reimbursement for direct costs of training. The beneficiaries will receive reimbursement for either 50% or 100% of all direct cost for any educational program that does not exceed two years in duration. It is true that the largest component of the cost of education is foregone income, which is not reimbursed; but it is precisely the role of the income maintenance payments to provide this element of subsidy for those families that benefit from combined income maintenance and manpower programs.

In order to measure the effect of cost reimbursement on training and work incentive, it is essential to ascertain that all people making these decisions possess the same information. Furthermore, if decisions are to be optimal, then they should possess all the relevant information that is available. From the experimental viewpoint, the absence of information for some of the families may introduce imperfections in the process of decision making and excessive variance in the initial conditions. In order to overcome these difficulties, we have introduced into the manpower component a function called "counseling." The goal of the counseling component is to offer to each eligible member (a person 16 years of age or older) of a participating family sufficient information about his opportunities and capacities to enable him to choose among alternative labor market options such that the choice will yield maximum satisfaction. The three components of the counseling function are as follows:

- (1) Assessment, in which each participant determines his labor market objective, identifies the barriers to achieving the objective, and develops a plan of action to reach the

objective. The counselor helps the individual identify realistic objectives.

- (2) Information services, in which the counselor provides the information relevant to each individual about present and expected future job market characteristics and available training and community services. In this component, the informational deficiencies are eliminated.
- (3) Follow-up, to ensure continued awareness of the program and to inform individuals of changing opportunities.

The information component will include help in making employer contacts and applying for jobs, but will not include direct job placement.

It is our hope that the initial counseling process will bring most participants closer to efficient utilization of their personal resources. Experimentally speaking, the initial counseling will try to achieve in a short time what a national program would probably attempt to achieve only over a longer period of time.

The cost reimbursement component of the manpower experiment changes the effective cost of increasing human capital, which is consistent with the approach to income maintenance, in which the negative taxes essentially act to change the effective price of leisure. The counseling function is introduced to compensate for market imperfections so that individuals can perceive the real cost and benefits of investments in human capital. The manpower program in our experiment provides a very simple treatment that will facilitate the eventual evaluation of elasticity of investment in human capital with respect to its cost. A manpower program comprised of actual training programs would have suffered from being too spe-

cialized and subjective because it would have been impossible to offer the range of training and educational options available from the existing system in Seattle and elsewhere. Thus, an econometric evaluation of such programs would not generalize.

Integration of Components

The experimental design provides for systematic integration of the manpower and financial treatment components described above. Table 1 summarizes the essential ingredients of the experiment. There are nine financial treatment options, comprising four tax systems for each of the two support levels, plus a null treat-

TABLE 1—TABLE OF DESIGN

	Support Level	Initial Tax		
		50%	70%	80%
No manpower	0			
	\$3,600	R=0 —	R=0 R=.025	— R=.025
	\$4,800	R=0 —	R=0 R=.025	— R=.025
Counseling only	0			
	\$3,600	R=0 —	R=0 —	— R=.025
	\$4,800	R=0 —	R=0 R=.025	— R=.025
Counseling and 50% subsidy	0			
	\$3,600	R=0 —	R=0 R=.025	— R=.025
	\$4,800	R=0 —	R=0 R=.025	— R=.025
Counseling and 100% subsidy	0			
	\$3,600	R=0 —	R=0 R=.025	— R=.025
	\$4,800	R=0 —	R=0 R=.025	— R=.025

ment. There are four manpower options comprising counseling, two levels of training subsidies, and a null treatment. Thus, in all, there are thirty-six possible combinations including a null control. It should be noted that not all possible financial treatment combinations are being used in the experiment. There will be no experiments with a 50% initial tax and a declining rate or with a constant 80% tax rate. The first is believed too generous, and the latter too confiscatory to be interesting as a program option.

The allocation of families to treatments will be accomplished with the aid of an "Assignment Model," which is an extension of a model used for assignment in the original New Jersey Income Maintenance experiment. The technical discussion of this model will appear elsewhere. We shall outline below some of the ingredients that were incorporated in the extended model.

The assignment model is designed to seek the allocation of families among treatments that minimizes some function of the error of the estimated response function subject to the budget constraint. To estimate the costs of alternative assignments, it was first necessary to hypothesize such questions as the expected attrition from the experiment, the expected use of manpower services, the expected reduction in work effort due to income maintenance payments, and how these will alter in the combined as against the separate treatments.

In addition to taking into consideration the size and income of the family in the assignment process, separate cell structures were created across two variables: race and number of family heads. It is believed that these two variables create potentially such different responses as to make it unwise to merge these into one experimental treatment. Thus the budget essentially considered allocation among

treatments for four separate experimental groups, as follows:

- (1) White families headed by one adult
- (2) White families headed by two adults
- (3) Black families headed by one adult
- (4) Black families headed by two adults

In this way, confounding effects of race or number of family heads could not occur in the experiment, and we would be able to recreate the experiment for each group separately. Thus, if responses are vastly different between these groups, the experiment would not fail on that account.

The Problem of Family Functioning

In addition to measuring work response of experimental families, the Seattle experiment will investigate the effects of the program on family functioning. The basic problem arises from the fact that income maintenance will provide additional economic security; at the same time the manpower program is creating changes in economic productivity of the family members. In combination, these changes may affect the relationship among family members and may either increase or decrease certain sociological responses by family members, such as changes in rate of desertion by fathers, changes in crime committed by juveniles, etc.

We expect to approach the problem as follows:

- (1) Incorporate in a thrice-annual interview various questions that will reveal some of the overall effects of the experiment on family life.
- (2) Select a subsample for in-depth study. No final decision has been made as yet on which tests should be conducted in the subsample al-

though we are considering forms of controlled family interaction experiments. We hope that our next report will shed more light on this question.

Some Problematic Issues

Being a totally new experience in the profession, income maintenance experimentation raises as many problems as excitement and hopes. We shall discuss below some of the issues that are fundamental to any controlled experimentation in economics.

Problems of Measurement

Concepts, such as income, wage rate, and number of hours worked, subsume very complex measurement issues. As we know, overcoming difficulties of measurement occupies much of the effort in any econometric research. In an experiment such as the Seattle Income Maintenance Experiment (SIME), one hopes to obtain considerably more accurate measurements and thus reduce measurement error to a minimum. In reality, however, reducing measurement error even in an experiment is difficult. Consider, for example, the care that must be exercised in defining terms such as "income"; distinctions must be made between transfer payments and income receipts, and between earned income and unearned income. This leads to very specific questions, such as, should income from property be taxed as unearned income or as earned income, representing the fruits of past labor? To resolve these issues, it has been necessary to devise complex "Rules of Operation" that incorporate the many individual decisions regarding definitions of family, definitions of eligibility, definitions of income, etc. Unfortunately, the rules of operation must incorporate a degree of arbitrary judgment.

The Hawthorne Effects

Hawthorne effects have become well known in social experimentation as being the experimental impacts on the family that are independent of the treatment itself. These need to be very carefully considered. If, for example, families on the income maintenance experiment develop a sense of "experimental responsibility," they may behave quite differently than a control group without such a sense of responsibility. We will attempt in the experiment not to communicate to the experimental group the nature of the dependent variables that are to be studied. In this way, their sense of experimental responsibility cannot be translated into an action directly relevant to the variables being measured.

The Issue of Credibility

Our purpose is to estimate the long-term response surface. This viewpoint is essential for the manpower component, which assumes that investments in human capital will require significant periods of time for a complete adjustment. However, the issue is "adjustment to what?". The families are told that the plans to which they are assigned will provide support for at least three years. But they know that the experiment will terminate. They also expect that some program will ultimately be established for all. These expectations influence in a deep way the long-term response of the families. Thus, if we provide a high support of, say \$4,800, and the families really believe that the ultimate support that will be established as a norm will be, say, \$2,800, then for the duration of the experiment they are likely to consider the \$2,000 difference as a transitory windfall, and we shall not succeed in estimating the true long-term response to a permanent support of \$4,800. Such issues are indeed so serious that all researchers

in this field must be extremely cautious in their evaluation of the results.

Unemployment in Seattle

During the design phase of the Seattle experiment, a situation arose that created a severe problem, the solution to which has greatly strengthened the overall experiment. In early 1969, when the decision was made by HEW to launch the income maintenance experiment in Seattle, unemployment in Seattle was about 3%, a rate below the existing national rate of 3.5%. Since that time, the situation has altered drastically and unemployment in Seattle is now over 10.5%, which compares poorly with a national rate of 5.5%, and is continuing to rise. This radical change in employment has made Seattle highly idiosyncratic; it is no longer possible to claim that the results of the experiment in Seattle will be representative of the nation as a whole, or even a significant sector of the nation.

The reasons a high level of unemployment in Seattle creates experimental problems are as follows:

- (1) Under conditions of forced unemployment the observed level of employment is not the desired level. Thus, our measurement of the hours, days, or weeks of work of the experimental population does not correspond to the level that would have been attained if employment opportunities were ample.
- (2) In the face of a high level of unemployment the private cost of training drops essentially to zero. Consequently people will tend to undertake an excessive amount of training and retraining, and the experimental measurement of the willingness to undertake training is vastly biased upward.
- (3) During periods of unemployment the degree of private risk increases significantly. Thus, even if the experimental families hold a job and want to adjust downward the number of working hours or days due to our treatment, they will probably hesitate because of the fear of losing the job altogether. Thus, the high level of risk may generate a "wait and see" attitude, and thus reduce sensitivity to treatments. More generally, the experiment seeks to measure changes in work effort due to treatment, and these may not be observable because human behavior will be idiosyncratic during periods of unemployment. During full employment, the great variety of job opportunities provides a low-risk environment where people may be more willing to experiment with changes in their life due to treatments.
- (4) High unemployment creates a downward bias in the measurement of changes in willingness to work. High unemployment creates a situation in which a person who wants to increase his work effort finds it extremely difficult to attain this goal; he finds much less difficulty in decreasing his work effort. Therefore you have, in general, a downward bias in measurement of the changes based on observed behavior. You will easily observe decreases in willingness, but increases cannot be translated into action. This is most important for the financial-manpower groups, where we hope to counteract negative effects of financial effects of financial treatments through the use of Manpower programs.

It may be argued, however, that all the

above argumentation is true both for controls and experimental groups. The answer to this is as follows:

- (a) With regard to Item 3 above, the degree of risk is so high that within the range of our support level there is not enough compensation for risk, and thus no differences in performance will be observed.
- (b) The argument may be true with regard to Items 1 and 2 above, if unemployment is uniform across all people in all industries, in all occupations, in all age groups, etc. Since we know that this is not the case, and unemployment has many points of concentration, it will be necessary to have an identical distribution of the controls in each experimental group. This is an incredible task, and the present size of the sample is much too small to enable the creation of equal distributions across so many variables. Furthermore, this task calls for the complete understanding of the variables that determine unemployment—a knowledge that is at present unavailable. Even more strongly, the dynamics of the economy may influence behavior in a way unrelated to preference.

However, the Seattle situation presented the experiment with a unique opportunity to provide national estimates of the effects of income maintenance for varying levels of unemployment. The decision has been made by HEW to launch an experiment of identical design in a city that has a socioeconomic structure and level of unemployment that mirrors urban America. In this way, answers can be obtained to the questions of program effects in the normal setting and in the setting of economic dislocation in Seattle, which in total will provide more significant policy guidelines than either of the parts done separately. We can learn which impacts are sensitive to the variations in the level of employment, and which are not. For those impacts that are affected by the level of employment, we may learn the extent of these program effects.

Status of the Experiment

As of this meeting, the experiment is being launched in Seattle. A total of 32,000 dwelling units were listed in the major low income areas of the city. On the basis of interviews conducted in approximately 25% of the city households, some 4,500 eligible families have been located and fully interviewed for baseline data. Approximately one-half of these will be enrolled in the program as either experimental or control families. Enrollment of these families is now underway. At a future time we will present a report on the sampling enrollment and interviewing and other procedures.

The Gary Income Maintenance Experiment: Plans and Progress

By TERENCE F. KELLY, *The Urban Institute*

and

LESLIE SINGER, *Indiana University*

The Gary Income Maintenance Project (G-X) parallels the New Jersey, rural, and Seattle efforts in that it focuses on the impact of alternative sets of income-maintenance structures on the work-leisure decision, consumption and investment patterns, and various aspects of family functioning. Gary was selected, however; because it contained a substantial population base not adequately treated in the other income-maintenance experiments; namely, black, female-headed families in a ghetto setting. Associated with experimental treatment of this rather unique group are special sets of conditions and problems. Our basic project design and analytical framework attempt to incorporate the uniqueness of the experimental setting.

Given the relatively large number of female-headed families in the sample,¹ there are at least three considerations which must be given special attention:

1. The importance of nonmarket work activity (cleaning, meal preparation, household management) in the labor-leisure decision process.
2. The role of welfare payments (AFDC) in establishing reservation wages.
3. The need to measure specific demands for the labor inputs of various, possibly noncompeting, demographic subsets of the population

¹ The sample is structured so that 60 percent of the families selected are headed by females.

(e.g., young males vs. young females).²

The next section of this paper is directed to the issues we felt should receive special attention as a result of the particular composition of our sample frame.

I. *Foci of the Gary Experiment*

The principal focus is on standard economic responses such as labor supply, consumption patterns and investments in human capital. In addition, we are concerned with the question of social services in Gary—in what amount they are demanded with and without income-maintenance transfers, and what effects they have.³ There is also a heavy sociological emphasis on such topics as family functioning, motivations and aspirations, but these will not be discussed further here.

Our primary concern, then, is the family work decision and how it is affected by the income and substitution effects of an income-maintenance transfer system. We intend to reconstruct and test the by now

² The question of measurement of specific demands for various types of labor is not addressed in this discussion. We plan an intensive periodic canvass of employers in the Gary area from which we will develop indices of job vacancies by type of occupation and by characteristic of employee demanded (e.g., male-female, level of education, etc.).

³ In terms of the impact of social services, the central hypothesis is that there is an interaction between income maintenance and services such that the combined effect of the two programs somehow differs from the sum of each program separately. In point of fact, we are unable to test the hypothesis adequately for all services. See below.

standard models, including the recent extensions proposed by Greenberg and Kusters, and Hall. This will involve the estimation of income effects and the income-compensated effects of changes in the wage rate (substitution effects) from what will surely be a rich micro-data set.⁴ Fortunately, our data base will allow us to give empirical attention to two extensions of the basic labor-supply model particularly relevant to the real world of Gary.

Our initial point of departure concerns the inclusion of indebtedness in the basic model. In this respect, we follow lines developed by Rosett and Lebergott.⁵ We hypothesize that an income-maintenance program will afford desired access to credit which was previously unavailable.⁶ An increase in indebtedness, in turn, produces a work-inducing income effect which partially offsets the direct leisure-inducing effect of the initial cash receipt from the program.

To the extent that these effects are operative, they require that we develop a two-stage analysis: the first stage showing the relation between the income-maintenance program and indebtedness (a relation which itself is composed of both de-

mand and supply elements) and the second stage showing the impact of indebtedness on the work-leisure tradeoff. Now, if we start from a position of equilibrium in the credit market, it is possible that the second stage will not become operative. A given family might increase borrowing in line with its increased income from the program—no more, no less.⁷ In this case, there will be no need for increased work effort to pay off newly acquired debts. If however, we allow for disequilibrium either initially (as is eminently plausible in the case of poverty-income families) or over time, the role of indebtedness is made considerably more important. To take an extreme case, suppose that there is some family-income threshold below which credit is totally unavailable and above which some positive amount is supplied by lenders.⁸ In this case, we are likely to begin with a situation of excess demand for credit which can be alleviated by the promise of a steady stream of riskless income-maintenance transfers. It is similarly possible that the family will desire (and be allowed) to obtain credit in an amount greater than the initial cash transfer which in turn will produce a labor-inducing (negative) income effect. We will attempt to test such hypotheses with data derived from a series of questions about indebtedness, not only to institutional lenders but to friends, relatives, and others as well.⁹

So far our model includes five essential elements: labor, leisure, current and fu-

⁴For example, we will be able to construct a more accurate set of potential wages than has previously been possible and to control better for specific demands than is generally possible with aggregate data bases.

⁵In discussing the Lebergott model, Griliches questioned the use of an indebtedness term, pointing out that both indebtedness and labor supply are affected by the same phenomenon—aggregate demand. The observed interrelation between the two, he argued, was to a certain extent spurious. However, Griliches' admonition is less applicable in a controlled experimental setting than it was in the macro-econometric model which he was addressing.

⁶A particular programmatic problem could develop here. Welfare payments are legally not subject to garnishment. The same will hold true for income-maintenance payments. It is therefore possible that such transfers will only marginally influence the willingness of lenders to extend credit to low income families. Myopia on the part of lenders excluded, there should still be an influence on the demand for credit which should lead to an increase in indebtedness by the treatment relative to control groups.

⁷We thus assume $d\text{Exp}/d\text{IM} \leq 1$, where Exp = total family expenditures (including the demand for goods purchased over time and thus reflecting desired indebtedness), and IM refers to the dollar transfer under the income-maintenance program.

⁸The illustration may not be so extreme after all, in view of the fact that credit ratings are often given in dichotomous terms.

⁹Another aspect of indebtedness with which we will be concerned is the period for which credit is received, a period which might well be influenced by the length of the experiment.

ture market goods consumption and asset accumulation. Yet another extension involves the dichotomization of the labor term into market and nonmarket elements.¹⁰

Specifically, we will expand our model to include market, nonmarket and leisure goods. With respect to the goods distinction, examples of those serving market activities are tools, uniforms (including white collars and requisite accoutrements) and some portion of the value of automobiles. The list of recreational goods is seemingly endless, from basketballs to boats. Nonmarket goods cover such items as vacuum cleaners, dishwashers and lawnmowers.¹¹ Discounting the differential prestige generally accorded Brand X as opposed to Brand Y, we assume that nonmarket goods do not enter the utility function directly—that their impact is on time saved, from which utility is derived.

The traditional basic theory predicts an unequivocal shift toward leisure away from labor as a result of both income and substitution effects associated with a negative-income-tax-type program. When we break labor into market and nonmarket elements, however, the picture is not so clear. In this regard, a number of observations seem appropriate:

a) The substitution effect, stemming from the fact that wages are taxed at a positive rate, operates only on market wages.¹² The reduction in hours spent in

market activity could be allocated to an increase in demand for both nonmarket and leisure activities, assuming neither to be an inferior good.

If market and nonmarket activities are substitutable, and if there are constraints placed upon the locus of market opportunities, leisure could come at the expense of nonmarket activity. This possibility hinges on two premises: first, that the family allocates a given amount of time to "sweat" (with both market and nonmarket activities being undifferentiably burdensome); and second, that there are likely to be some institutional constraints placed on hours of market activity (such as the standard work week). If these constraints are severe enough, the entire shift to leisure will likely come from nonmarket activity. To the extent that constraints are less rigid or that the two activities are imperfect substitutes (or are complementary), more of the shift to leisure will be borne by market activity.

b) In the simplest case, we might expect the income effect to be entirely leisure producing, with nonmarket activity remaining constant.¹³ However, a number of highly probable purchases are likely to be substitutable for work in the home (vacuum cleaners, dishwashers, lawnmowers, to say nothing of maids,¹⁴ etc.).¹⁵ Here the income effect causes a reduction in nonmarket activity, so that in the extreme there is a multiple impact which is

¹⁰ There is certainly nothing new in this dichotomization. While traditional models mention the importance of nonmarket activities, data constraints have prevented the mention from being more than mere lip service. One particularly ingenious exception is found in Altman and Barro. We will attempt to give an empirical base to those authors' theory.

¹¹ As should be obvious, we are following the Lancaster approach in considering characteristics of goods. Note also that the performance of nonmarket activities is not limited to females, in contrast to most previous literature.

¹² In G-X some tax rates for women previously on AFDC will actually be reduced under the income-

maintenance program. In this case, the converse (i.e., a shift to market labor) applies.

¹³ That is, if we view the income effect in terms of time, rather than goods.

¹⁴ The maid example is not really so far-fetched if we consider care of children as part of home work activity. In Gary a similar service, day care, is offered to a portion of the sample.

¹⁵ We do not give full justice to the complexity of the problem. To the extent that the income effect expands consumption horizons and/or that market time and nonmarket goods are complementary, the income effect is larger than in the simple static case where nonmarket goods and market time are substitutable.

leisure inducing. But the chain of events does not stop with the first round. As time is diverted from nonmarket activity, it can be allocated to market activity as well as to leisure. In this event, the second-round impact could serve to diminish the original impetus to reduce market work. This tie between consumption goods and labor activity, while recognized in the literature, has not been accorded the empirical investigation it deserves.

c) According prominence to nonmarket activities requires that we give further thought to the interdependence of family labor supply decisions. Certain chores in the home are traditionally performed by women, others by children. To the extent that the supply of labor to the locus of available work hours in various activities is less than continuous and differs for alternative family members, the traditional marginal model is inappropriate. If women can enter the labor market at less intensive jobs than men, an income-maintenance program could cause a reallocation of market work in favor of women.¹⁶

Work in the home is often liberally interspersed with leisure. And the same problem also exists in the case of market and leisure activities (extended coffee breaks and general loafing, for example). Clearly we are not at a stage where we can measure the allocation of time and utility across leisure, market, and nonmarket activities. We do hope, however, to be able to construct indicators of *change* in non-market activities. These indicators will be constructed from two sources. First, we will ask direct questions about the amount of time spent in various home chores (e.g., about how much time do you spend doing

the laundry, cooking, etc.?) and in leisure activities (how many hours do you spend reading, watching television, etc.?).¹⁷ Second, we will observe the purchase of certain items (vacuum cleaners, dishwashers, other appliances) which appear to shift the home production function in favor of less labor-intensive processes. Until knowledge about this rather hazy area is expanded, we will have to content ourselves with this admittedly crude first approximation.

The Gary project was specifically charged with the responsibility of investigating the demand for social services¹⁸ (such as day care, family counseling, homemaker services, information about access to housing, etc.) under alternative family income conditions, because there is lively policy interest in the optimal relationship between such services and income-maintenance programs. This charge is reflected in the sample design as follows.

There are two central issues involved. The first concerns the tradeoff between the demand for cash transfers and that for services. What is the residual demand for services once basic cash needs are approached? Second, is there any interaction between the receipt of services and of income-maintenance transfers such that the social benefits deriving from the two programs in combination exceeds the sum of the benefits when each program is operated in isolation? To the extent that there is a positive interaction, results yielded from an examination of the two separate programs could suggest a suboptimal level of social investment in both programs.

¹⁶ What is being suggested here is that the family labor supply decision should be viewed as a linear programming problem. Alternative family members can engage in any of three activities with which are associated alternative (direct and indirect) utilities. The utilities associated with each activity vary across the respective family members.

¹⁷ We are aware of the measurement difficulties here. Eating out involves food consumption which would otherwise require home preparation, but it can also involve the purchase of status or simply a good feeling of being pampered.

¹⁸ The social service treatment is an important and complex portion of the Gary experiment. Space does not allow full discussion of the question.

Unfortunately, from an experimental point of view the measurement of the two phenomena requires conflicting strategies. If an unconstrained market demand is allowed to develop, self-selection might bias the observed impacts (the benefits) of services. At the same time, if individuals are experimentally assigned to various service treatments, there is no way to measure the demand which would obtain in a free choice market.

We have decided to pursue a mixed course of action. For day-care services we will assign families to treatment cells, where treatment in this case refers to varying levels of day-care subsidies ranging in value from zero to 100 percent. In that way, we will measure both the demand for day care and its impact on labor supply and family functioning, alone and in conjunction with income-maintenance transfers.¹⁹

In the case of other "softer" services, we will investigate the unconstrained demand question. Virtually nothing is known about this area. In the first year of project operation, we will simply determine a relevant mix of services. This will consist of informing the treatment group of alternatives in an attempt to create a situation approaching perfect information. Service utilization patterns will be measured and questions will be asked about those services which sample families feel meaningful and desirable. In the second and third years of operation we will attempt to introduce fees in order to measure both price and income elasticities. While it would also be interesting to measure the *impact* of services, we feel that there is insufficient knowledge to permit us to identify those variables on which

services might operate.²⁰ Consequently, we are beginning with the first step of the two-stage question.

II. *Experimental Design*

The Gary experimental design calls for the creation of twelve distinct cells by completely crossing three main independent variables and partially crossing a fourth:

- a) Income-guarantee levels—two experimental levels of \$4,200 and \$2,700 (for a 4-person family) plus a control group which receives no payments.
- b) Sex of household head.
- c) Place of residence, specified in terms of whether the family resides inside or outside the Model City area. It is necessary to make this distinction because Model City residents have access to a number of social services which are not available to those living outside. The most important of these services is complete access to subsidized day care to which Model City families will be randomly assigned.
- d) Day-care subsidies, offered to Model City residents.

The total sample of approximately 1,600 families is to be distributed among the 12 cells as shown in Table 1. A family is defined here as an adult and at least one dependent child. The total sample is evenly split between guarantee levels, split 13-7 between Model City and non-Model City, split 60-40 by sex of head of household. The choice of the particular variables and the proportions assigned

¹⁹ One of our central hypotheses is that day care shifts the family production possibility curve outward, thus mitigating the potential work disincentive effects associated with the cash transfers.

²⁰ We should note that we contracted a number of leading professionals in the social work field and asked them to list just what they thought services were supposed to do. In no case did we receive a list of measurable outputs. This attests to the difficulty of the problem.

TABLE 1—BASIC SAMPLE DESIGN

	Inside Model City (Day Care)			Outside Model City (No Day Care)		
Income Guarantee Level	\$4,200	\$2,700	\$0	\$4,200	\$2,700	\$0
Sex of Head:						
Male	104	104	208	56	56	112
Female	156	156	312	84	84	168

N = 1600

Note: Figures are approximations. The sample is allocated by the Watts-Conlisk allocation model. We have not made the final allocation run at this time.

across the various cells was based on a variety of considerations.

The lower guarantee level had to be set to dominate payments by the County Welfare Office in order to induce participation in our project. Limited funds necessitated that the level not be set too high, however. The upper guarantee level of \$4,200 was chosen to maximize expected variance in the income effect, subject to the budget constraint. In addition, we wanted to test a level which might be of interest to the policymakers in the relevant future. Four or five years from now, a \$4,200 level might well be the subject of considerable discussion.²¹

Our interest in assuring a predominance of female-headed families stems from the fact that the policymakers feel that this group has received insufficient attention from the other experiments.

Because the Model City subsample is subject to additional treatments (the availability of both day-care and social services), and because there is likely to be a certain amount of self-selection of these treatments, we feel that more cases will be required for the Model City cells. Further, the Model City more closely approximates the ghetto conditions which the ex-

periment is designed to reflect than the outlying areas.

The day-care portion of the experiment is particularly complicated. Subsidies of varying amounts (tentatively 33 percent, 67 percent and 100 percent)²² are offered to all families (including controls) residing in the Model City. In addition, these families are guaranteed slots in the day-care program should they wish to avail themselves of the opportunity. Families living outside the Model City serve as controls. They receive neither subsidies nor guaranteed access to day-care slots. Our design thus allows us to measure the demand for single-priced day care under alternative income conditions, as well as the impact of income maintenance alone, income maintenance in conjunction with the opportunity for day care, and day care alone on, say, labor supply.

The design was dictated by the administrative and political realities of Gary. It was not feasible to offer subsidies and the assurance of day-care slots to one portion of Model City residents while excluding another portion. In contrast to the other experiments, we use but one tax rate—50

²¹ Naturally, both guarantee levels are scaled for alternative family sizes and will be adjusted for changes in costs of living.

²² For the first and subsequent children. A family wishing to place two children in day care pay twice the amount which would be paid for the first, and so on. Since this is a study of demand, rather than costs, we felt it unadvisable to attempt to allow for economies of scale.

percent.²³ The choice of a single rate was a compromise. Initial interviews convinced us that respondents had a great deal of difficulty in understanding the basic tenets of a negative income tax program. Maximum comprehension required minimum complexity. This is particularly true for the sociological portions of the experiment, which have high priority in Gary. Further, since a substantial number of the experimental families will be former AFDC recipients, a comparison of pre- and post-experimental tax rates will, in any case, yield a wide array of tax rate changes. Effective tax rates on welfare grants are extremely complicated, depending in part on allowable deductions, which in turn are often left to individual case worker discretion. With the use of welfare records, we will create case histories for each of our former AFDC respondents and attempt to construct effective tax rates. Labor responses will then be related to *changes* in effective rates, or to a comparison of differences between experimental and control groups.

At one time we considered introducing accounting periods as an additional classificatory variable. We have since decided against this. The bulk of the cases will be given a single accounting period (four weeks), with a few subjects randomly selected from each of the eight noncontrol cells to participate under a separate accounting period (12 weeks). Six to ten cases will be selected from each of the four Model City cells and three to five from those of the non-Model City cells (thereby not disturbing the relative cell frequencies). By selecting a small number of subjects from each noncontrol cell for participation under the alternative accounting schedule, comparisons can be made between those selected and all other

subjects in order to test the null hypothesis of no differences between accounting-period groups.

III. AFDC in Gary

Some preliminary work has been done on a sample of 792 case records of AFDC families. All of the families have been on the welfare rolls for at least five years and were in the labor force for at least three years. A substantial part of the sample had been on welfare for more than ten years. These families experienced a substantial change in the AFDC program when, in the second quarter of 1967, higher AFDC grants, lower tax rates on earned income, and other liberalizations were instituted.

Table 2 shows the "permanent value" of several crucial variables for the pre-1967 period and for the post-1967 period. The results are shown separately for the long-term AFDC families (on rolls for at least ten years) and for the medium-term ones (five years or more). The "permanent values" were formed by using an exponentially weighted moving forecast formula and taking the last value of the forecast series.²⁴ Data from 1960 to 1966 were used for the pre-'67 projections and from 1967 to 1970 for the post '67 data.

As will be noted from the table, the liberalization of the welfare laws was associated with an increase in wage income. For long-term AFDC families it was also associated with a small amount of work reduction which was more than compensated for by wage increases. The effects of the liberalization are quite evident in terms of lower tax rates and higher benefit levels. The long-term cases seem to be characterized by lower wages, higher tax rates, less

²³ This is the rate at which the grant is reduced as other income rises.

²⁴ For more information and derivations see: Leslie Singer, *Optimum Information Predictors*, Chicago, 1969, pp. 71-81; Leslie Singer: "Quasi Families, Poverty and the Reluctant Acquisitor," unpublished discussion paper, Indiana University.

TABLE 2—"PERMANENT" VALUE OF SELECTED VARIABLES FROM AFDC CASE RECORDS

	Medium-term Families (5 yrs. or more)		Long-term Families (10 yrs. or more)	
	Pre '67	Post '67	Pre '67	Post '67
Wage Income/mo.*	\$69.91	\$111.52	\$49.11	\$ 59.25
Hourly Wage Rate*	\$.70	\$.89	\$.65	\$.75
Hours/mo.	112	133	64	60
Intensity of Work (function of full-time)	.31	.49	.40	.39
Potential AFDC Grant*	\$71.20	\$ 89.34	\$93.67	\$155.16
Implicit Average Tax rate	.25	.07	.19	.17
Implicit Marginal Tax rate	.36	.20	.57	.27

* Dollar values have been deflated to a 1957-59 basis.

work and greater potential benefits. The latter may be due to larger family sizes.

These crude calculations both indicate that there is a nontrivial amount of work activity among even the long-term welfare cases, and also strongly suggest that such activity is sensitive to the amount and terms of the cash transfer program. They underscore the importance of including the female-headed welfare family in the over-all program of experimentation with alternative forms of income maintenance.

IV. *Non-Linear Alternative*

Besides the conventional linear treatments (two alternative guarantee levels and one tax rate) consideration is being given to use of a more complex formula for a small sub-sample of the families. All the experiments to date have used a benefit formula which is based on family income (somehow defined) and family size. It may be possible to devise formulae that are less damaging to incentive for working and hence more efficient by making benefits depend as well on hours of work and/or wage rates.

When benefits are based on income only there is no distinction between the person who earns \$50 a week at a full-time job and one who has twice the earning power but chooses to work only half time. It

might be desirable to make such a distinction and pay a larger benefit to the person who is making the greater effort at self-support. One way of accomplishing this would be to get reports on hours as well as income and make the guarantee a function of hours. For example, anyone who works half time or less is eligible for only 80 percent of the basic yearly guarantee (\$2700 or \$4200) which is then reduced by a fraction of the amount of income earned (say 50 percent). Those who work full time or more would be entitled to the full guarantee (similarly reduced with income) and those between full and half time would receive a scheduled proportion between 80 and 100 percent.

One could also adjust the marginal tax rate according to average wages so as to have minimal "price" effects on those who earn very low wages (this might imply a high break-even, but no one could work enough hours to reach it). The tax rate would be relatively high for high-wage persons, making their break-even point low and this would prevent wasting a large amount of benefits on those who presumably don't need them. So long as the high wage persons continue to work, the high marginal rate does not affect them because they are above the break-even point.

One or both of these might be a useful feature for improving the efficiency of income-maintenance programs. They would require getting reports on wages and/or hours as well as income, and the feasibility of this needs to be assessed. With such features it would be possible to add some further incentives to the system. For example, education or manpower training could be counted as hours worked—as could, perhaps, some part of child care which enabled someone else to work. Depending on the way that such activities entered the formula for computing average wage rates, they could have an effect on the effective tax rate as well.

Further analytic work is needed to determine the most promising ranges over which the tax rate might be graduated.²⁵ There was, in any case, an evident interest in a variety of schemes, including wage subsidies, in the recent debate over the

Family Assistance Plan and it would seem quite desirable to have some experimental evidence on a wider class of formulae than have been used so far.

V. A Final Word

No experiment paper should be complete without mention of possible response bias. Clearly the experiments described today may well encounter response biases stemming from several sources. The Hawthorne effect and the transitory nature of the experiment are but two examples. The Gary experiment includes a scheme for trying to measure such bias which is absent from all the others. The bulk of our sample is being interviewed three times a year. As a first step in accounting for any biases, we will select three additional control groups. The first group will be interviewed only once a year, the second will be interviewed only at the beginning and end of the experiment, and the third will be interviewed at the end only. This procedure should allow us to estimate the testing effect upon our participating families.

²⁵ For details of the plan and other considerations see: Leslie Singer, "Metropolitan Area Roundup," *Indiana Business Review*, Sept.-Oct., 1970: 45; Leslie Singer, "The Gary Experiment in Income Maintenance," unpublished discussion paper, Indiana University.

DISCUSSION

JAMES N. MORGAN: Economics is rapidly becoming a science with three major branches instead of one. There is the traditional and powerful analytic science working out the interrelationships within systems: resource allocation, pricing, dynamic equilibrium. There is the new behavioral science of economics, finding more precise information on the way people actually make economic decisions so that we can better specify the parameters of the analytical models. The experiments we discuss

trolled charges or differences that it is difficult to isolate the impact of any one explanatory variable. The experiments under discussion fall somewhere between these extremes and may hopefully avoid the worst problems of each.

To come to the problem at hand, namely, the possible effects of different "rules of the game" on people's work behavior and their family planning behavior, it is useful to state a general theoretical model:

Change in the rules	→	Awareness → and Credibility	→	Interpretation and attitude response	→	Behavioral changes	→	Resulting changes in earnings or birth rate
------------------------	---	-----------------------------------	---	--	---	-----------------------	---	---

today reflect, hopefully, a great leap forward in economics as a behavioral science. A third area, mostly only looming on the horizon under the title of social indicators, is the assessment and measurement of what matters to what subgroups of people—the social welfare function, if you will. As we make more and more decisions on resource allocation politically, our elected representatives need better information on what matters, how much, and to whom. The welfare, that we were all taught could not be measured, at least must be compared for competing alternatives.

To return to the second area, much of our information on economic behavior has come from analysis of time series, progressively disaggregated but limited in scope and by their very aggregation. It has also come from survey data, cross section and reinterview panels, providing some "natural" or quasi-experiments, but not always the ones we want most. If we are impatient, the alternative seems to be experiments.

The term "experiment" has been used to cover a range of things from rigidly controlled, somewhat artificial, laboratory experiments, producing rather precise information on very restricted questions, all the way to public information or persuasion campaigns in selected areas. Those at the first extreme are sometimes difficult to extrapolate to the real world, where people must continue to work together, while those at the second extreme are frequently so muddled up by other uncon-

Messing this up are other changes in the rules, or other changes in the situation, which can affect this causal sequence at any stage (and of course the subsequent ones). Attitudes or behavior can even change from sheer self-consciousness that someone is watching (the Hawthorne effect). The rules change, differentially from place to place and even as between people, or their administration (particularly by local welfare officials and social workers) can change. Wage rates can change, so that earnings could stay constant with less work. And, of course, the local employment situation and job opportunities can change.

As a general rule, unless one can be surer of controlling these exogenous changes, the presumption is in favor of widely dispersed experiments so that randomization and statistical controls can take care of the problem. I have never understood the notion that one should conduct negative income tax experiments in limited localities, and still think it is a mistake. It is doubtful that enough recipients know one another so that one could get the contagion effects of a real program.

But a more important mistake would be to focus on the mechanics of the rules (the levels of support and the marginal tax rates) at one end, and the ultimate outcome in work or pregnancy at the other, to the neglect of the intervening variables. There are two main reasons why this is a mistake. First, the intermediate changes occur earlier in time and, second, they are less subject to the distorting in-

fluences of outside events. It is possible for a family given better economic security to start using family planning and to look harder for work, but the first methods they try may not work, or the job market may deteriorate for their particular skills or location. And change in work effort will not correlate perfectly with change in earning, so a finding that their *behavior* has changed is really more convincing than the finding that they are earning more, or less, or having children or not. The "Seattle problem" of growing local unemployment provides a dramatic demonstration of this possibility.

Why, you may ask, should we be concerned with "awareness," or information and understanding, or attitudes, or evaluation? Partly because, as the Watts paper shows, it is never a sure thing that people know the rules we think they are responding to, and partly because there may be things about the other rules they face which we ourselves do not understand. And they must interpret what it all means in terms of paths to their goals.

Indeed, another problem in all these studies is just what the other rules are. For instance, for those not on welfare, the Federal food stamp plan has a built-in 30% marginal tax rate. That is, if the monthly income goes up by \$100, the *absolute* size of the food stamp subsidy drops by \$30. Those in subsidized housing may or may not face another tax—a rise in rent with income—often close to 20% although it may be a jumpy schedule. What do people *think* will happen to their subsidized child care, or the availability of free medical care, if they earn more? What about extended family responsibilities?

Failure to think in broad motivational terms leads to such absurd concerns as that with the sharp drop in marginal tax rates when the subsidy ends and the family starts *paying* taxes—usually a drop from 50% to 14%. What is wrong with that? Indeed, considering the other benefits one loses by becoming economically independent, such a lure may be very important. Attempting to reduce the tax rates as earnings rise, at the expense of having them start high, could be the worst possible motivational device. If inertia and disbelief in the rules are important, one could

argue for zero or very low tax rates for the first earnings (also reducing enforcement problems) and a gradually rising rate.

Finally, there is the important fact that the constraints under which people operate are often rigid, and usually not additive. New analysis techniques to deal with such problems must be employed in place of the old favorite additive regression models. It is the activity at the margin which interests us, and many people are not at the margin—that is, they really have no choices to make. In the classic words of James Duesenberry:¹ "The difference between economics and sociology is very simple. Economics is all about how people make choices. Sociology is all about why they don't have any choices to make."

There also is simple inertia, leading to no response until some threshold is crossed. Once choices are being made, social psychology has something to contribute because it is concerned with why people make the choices they do in the situations where they have some freedom.

There is also a serious problem in these studies with the short run and the long. I have a hypothesis that the major advantage of providing basic economic security for people is that it reduces fear and the resulting regression or aggression, narrow horizons, and unadaptive behavior. The result of providing security might well be the unleashing of energy, enterprise, imagination, and planning no one believed possible from such people. But this takes time and a belief that the security is there—i.e., that the guarantees will last. The present experiments throw little light on this and seem more likely to test another hypothesis, namely, that one can get people "hooked on money," so that the ending of the program will have more dramatic behavioral results than the program itself. But those behavioral results may well be frantic, maladaptive, uncreative, and fruitless. Some follow-up a year or so later would seem to be called for to investigate this.

And if people really think of the present

¹ James Duesenberry, "Comment," in *Demographic and Economic Change in Developed Countries*, Princeton, 1960, p. 233.

scheme as a short-lived one, they might well (rationally, most of us would say) put much of the money in consumer durables, so they would have something to show for it when the period is over.

Given all these problems, my immediate suggestion for these programs is that they concern themselves with measuring more of the intervening variables in the process, so that we know more about where the chain breaks down, if it does, or what else happened to alter the outcomes, and that they measure the varying constraints and changing environments that people face, and whether they are *aware* of them too.

I have a few small comments or queries.

On the paper by Terence Kelley and Leslie Singer:

Do we know something about the order of magnitude of some effects? For instance, would the labor hours saved by household appliances really be large enough to produce any visible effect on the labor supply?

Is the amount of day-care subsidy as important as the sheer availability of any day-care facilities that are convenient, or the feasible solution of a transportation problem?

Do people at these levels use much credit anyway, and, even if they could increase debt on the basis of non-garnisheeable payments, would this be likely to affect their behavior *per se*?

Are not constraints more important anyway: disabilities, need to care for other disabled family members, lack of available jobs, no feasible child care facilities (too many children)?

Why the concern for leisure, as distinct from sleep and maintenance time? Non-market work, yes, but that should be enough.

The notion of altering benefits according to the amount of work done sounds fine and fits with our notion that welfare should be measured by a function which includes a term for the free time available to enjoy the fruits of one's labor. But what about the intensity of the labor? A night watchman or cab driver job as a second job does not compare with two laboring jobs, surely.

The complex design that is still not completely factorial, but partially confounded,

will leave one unsure whether the day-care effects are really Model City effects, or whether the effects elsewhere might not be different.

On the paper by Robert Spiegelman and Mordecai Kurz:

The arguments for a declining marginal tax are unconvincing. A much stronger case could be made for a low or zero starting rate to encourage the first entrance into the labor market, and higher rates later when the goal of independence is close and the loss of certain other "benefits" looms. The free medical care of welfare and other subsidized programs with means tests all tend to be lost "at the margin," so a precipitous drop in the marginal tax rate to 14% is nothing to worry about.

Indeed, the combinations confound things so that one cannot really compare the 50% and the 80% rates unless certain other effects turn out to be unimportant. And if understanding is important, the counseling may seem to have an effect here, when it would not be needed with simpler programs. There is also a problem how one can measure the *quality* of the counseling input. This is a problem that has plagued many demonstration projects.

On the paper by Harold Watts:

I am impressed by the very low loss rates in the New Jersey experiment. However, there must also have been some initial losses, people who were not gotten through the enrollment interview, and some potential biases might arise from that. I doubt that they are important, but the overall assessment should start with a representative sample of the eligible families (including those missed even on the screening interview).

The low level of understanding of the rules demonstrated in New Jersey makes the advisability of the complex assortment of treatments being used elsewhere doubtful. More important, it makes it crucial that in the later interviewing some attempt be made to find out what people tried to do and why, in case other misunderstandings continued.

I should point out that the experimental designs reflect a good deal of sophistication and that there seems to be an appropriate concern with measuring and accounting for the

changes in the environments of the experimental families other than those produced by the experiment. There is, however, a possibility that building so many things into the experimental designs, together with some necessary and sensible confounding of an otherwise pure factorial design, may well turn out to make it impossible to find much that is statistically significant. This is a common problem with experimental designs—building in too many dimensions relative to the amount of replication—and becomes a problem if there are not only main effects but a number of important interactions as well, or if there are uncontrolled (exogenous) changes correlated with

the manipulated variables. If there are some extremely important outside events which affect people, then the final analysis will have to be a kind of analysis of covariance, using up a great many degrees of freedom.

None of what I have said should take away from the fact that we are witnessing an important event in the development of economics as a behavioral science. I feel a little like those who were called upon to evaluate the Kinsey Report. Whatever the shortcomings, whatever my quibbles, the importance of the undertaking is such that we should all wish it well and try to make our comments as constructive as possible.

THE STATE OF ECONOMICS: THE BEHAVIORAL AND SOCIAL SCIENCES SURVEY

Introductory Remarks

By HENRY W. RIECKEN
Social Science Research Council

The subject of today's session is a report by the Economics Panel of the Behavioral and Social Sciences Survey (BASS). This report is one in a series issued under the general sponsorship of the Behavioral and Social Sciences Survey Committee, a group organized under the auspices of the National Academy of Sciences and the Social Science Research Council to review and report upon the various behavioral and social scientific disciplines.

This activity had its origin in the program of the NAS/NRC Committee on Science and Public Policy (COSPUP) which had been established by the Academy in response to a request from the Congress for advice on policy for the nurture and development of all of the sciences. The response of COSPUP was to commission a series of reviews of various scientific disciplines, beginning with physics and chemistry and proceeding through astronomy, biology, mathematics, computation, and other fields. The objective of these surveys of scientific fields was to provide advice to administrators on the needs and opportunities for the development of these several disciplines and to make a contribution to the formulation of a rational policy for science. The audience for these reports was intended to be primarily the Congress, but also members of administrative agencies in the federal and state governments, university adminis-

trative officers, the professional staff of private foundations, and others who were concerned with administrative decisions that might affect the growth and development of scientific disciplines.

When confronting the social and behavioral sciences, the Academy turned to the Social Science Research Council and suggested the collaborative establishment of the BASS Committee. This committee was organized in 1966 under the general chairmanship of Ernest R. Hilgard of Stanford and Henry W. Riecken of the Social Science Research Council. Its principal professional staff was Stephen Viederman.

The basic organization of the survey was in terms of academic disciplines: anthropology, economics, history, geography, political science, psychology, psychiatry, sociology, and finally statistics-mathematics-computation. The chairman and cochairman of each of these disciplinary panels, together with the overall chairman, constituted the "central planning committee" of the survey. That committee was constituted in the summer of 1966 and the individual disciplinary panels began work that fall. Each panel was charged with the assessment of its own field and recommendations regarding its future development. The panelists were asked to describe in terms comprehensible to an educated layman the present state of the field, its achievements and contributions, its prospects and promising areas of

future development, and finally its manpower and financial resources and needs in the future.

The survey was financed by grants from the National Science Foundation, the National Institutes of Health, and the Russell Sage Foundation. The effort lasted a little more than three years with the first product being an overall report entitled *The Behavioral and Social Sciences: Outlook and Needs*, issued in November 1969. This report covers all of the social science and behavioral disciplines and reports in considerable detail the outcome of a survey of Ph.D. granting departments in each of the disciplines with which the survey was concerned. The survey covered manpower, research financing, space, projected enrollments, degrees granted, space and equipment needs, and similar matters. Projections through 1972 were also derived from the survey for most of these items.

In addition to the overall report, nine disciplinary panels in the fields mentioned above have been published. The report to

be discussed today is that of the economics panel which consisted of Carl Kayser and Robert Solow, cochairmen, and Hollis Chenery, Arnold Harberger, Dale Jorgenson, James Morgan, Guy Orcutt, Joseph Pechman, Melvin Reder, Nancy Ruggles Henri Theil, and James Tobin. The report consists of chapters describing subfields of economics, written by the individual members of the panel, together with additional contributions from Gardner Ackley, Richard Caves, Jerome Rothenberg, and Vernon Ruttan. Nancy Ruggles was responsible for editing the several contributions as well as for the two chapters on "The Goals and Achievements of Economics" and "The Organization of the Profession and of Research in Economics." The final chapter of the report is a summary of recommendations by the panel for the future support and development of economics.

The panel report on economics, as well as the overall report, has been published in hardback and paper cover by Prentice Hall.

The Reviewers Reviewed

By CHARLES L. SCHULTZE

Brookings Institution and the University of Maryland

The *Economics* report of the Behavioral and Social Sciences Survey is a "mixed bag," to lapse into the current jargon. It is a part of the overall survey of the behavioral and social sciences, jointly commissioned by the National Academy of Sciences and the Social Science Research Council in 1967. The purpose of the survey was to appraise the current status of knowledge in the social sciences and to make recommendations for national policy aimed at further developing these disciplines. A panel of scholars in each discipline was appointed to review and evaluate its own field. An overall report was published in early 1970,¹ followed by publication of reports from the individual panels for each discipline.

The *Economics* report,² consists of a series of chapters covering fourteen major subspecialties each written by a well-known member of the profession. Two introductory chapters and a summary of recommendations were contributed by Nancy Ruggles, who also served as overall editor for the volume.

The report of the economics panel of the BASS survey can be evaluated at several levels. Viewed as a summary of the current status of economics as a technical discipline, the report is a highly competent piece of work. Each of the fourteen authors has described the progress of the discipline in his specialty—succinctly, carefully, and on the whole with judicious choice of what to emphasize and what to

ignore. Recommendations for needed areas of further research flow from each author's intimate knowledge of his specialty, and taken altogether provide a detailed menu for future technical work. Given the eminence of the authors one could hardly have expected less.

Yet one comes away from the report with a curious lack of intellectual excitement. With some exceptions—particularly the essays by Guy Orcutt, "Simulation, Modeling, and Data"; James Morgan, "Survey Research and Economic Behavior"; and Jerome Rothenberg, "Urban Economics"—the report does not convey a sense of excitement, either about new discoveries looming over the horizon or challenges soon to be tackled. No double helix in embryo lies in these pages; no miracle rice is about to be bred. While the report is not smug, as Professor Gurley would have us believe, neither is it alive with the awareness of great problems to be faced and new worlds to be conquered. Nancy Ruggles' closing sentence in her introductory chapter, "The Goals and Achievements of Economics," reads:

"Although the past achievements of economics have been substantial, they may well be dwarfed by the unsolved problems that face complex modern societies."

However valid this conclusion, one would scarcely reach it by reading the individual chapters.

Unlike the situation in the hard sciences, the frontiers of economic research are only in part dictated by the internal logic of the discipline. They arise perhaps even more directly out of those emerging social concerns which dominate the politi-

¹ National Academy of Sciences and Social Science Research Council *The Behavioral and Social Sciences Survey*, Englewood Cliffs, New Jersey, 1969.

² Nancy D. Ruggles, editor, *Economics*, Englewood Cliffs, N.J., 1970.

cal discussion of the time. And in recent years, for good or for ill, this tendency has been strengthened by the attitudes of many of the younger members of the profession who have become committed, in growing proportions, to use the tools of the discipline in analyzing and solving the problems of race, poverty and environmental pollution. From the radical left, the profession has come under mounting attack on grounds that its habits of thought and its very tools of analysis are so conditioned by the status quo and the capitalist institutions of Western society that it is incapable of addressing the critical problems of our day.

I should like to present still another view: *first*, that far from being irrelevant, the basic analytical tools of our profession are peculiarly well suited to grapple with many of the central problems of modern society; *second*, that in many ways the profession has missed the boat in perceiving how its analytical tools can come to grips with these problems; and *third*, and paradoxically, that the more we turn our attention to current social issues, the more we shall have to adopt some of the methods and approaches of the hard sciences.

Let me turn to my first point. There are several characteristics of modern society and the problems it faces which lend themselves particularly well to analysis and attack by the tools of our trade.

In the first place, the nature of the social problems with which public policy deals has changed sharply in the last generation. During the 1930's and early post-war years the major themes of social policy were three-fold: (1) correct the consequences of economic instability by overall monetary and fiscal policy, leaving basically unchanged the behavioral and institutional structure of society; (2) regulate those areas where private behavior seemed to lead to unwanted consequences—hence, the SEC, the NLRB, etc.; (3)

provide income supplements to those *unable* to work (the aged and the disabled) and temporary supplements to the *experienced* unemployed.

While the problems dealt with by those measures are still a matter of concern, the cutting edge of much of recent public concern and some of recent public policy lies in measures explicitly designed to change social behavior and social institutions. In stabilization policy we are indeed still worried about how to reach a particular target point on the Phillips curve (whence, concern about lags, monetarists vs. fiscalists, etc.). But our attention has also been caught up with how to shift the Phillips curve, through attacking problems of structural unemployment, discrimination in the labor market, obstacles to labor mobility, and so forth.⁸ A list of other major national concerns would also exhibit the same emphasis on changing social institutions or social behavior: compensatory education to redress past imbalances, work incentives in income maintenance programs, restructuring the health delivery system to make it more effective and to channel more of its services to the poor, reducing environmental pollution, improving the effectiveness and the equity of our criminal justice system, and so on down the list.

One common thread runs through all of these concerns and the attempts to deal with them. They require joint action by a host of decision makers, public and private. While public programs which produce purely public goods (like national defense), or simple income distribution (like veterans pensions) are still important, the major thrust of current social policy deals with matters in which a hierarchical command and control system can

⁸To avoid unneeded controversy with believers in the "natural rate" hypothesis, this sentence can obviously be reworded in terms of reducing the "natural" rate of unemployment.

no longer work. Improving the health delivery system, reducing structural unemployment, and controlling air and water pollution are not areas in which centralized decision makers can alone determine results. Even where programs are predominantly confined to the public sector—as is the case with elementary and secondary education—the major instrument of national policy is not direct action but joint action with state and local governments through the grant-in-aid mechanism.

To a growing extent, therefore, the solution of major social problems depends upon the behavior of a large number of *independent* decision makers, public and private. National objectives must be reached in thousands of diverse communities across the land. There is no hierarchy of officials in a single line of command which can be directed toward a set of predetermined objectives. A highly centralized managerial system cannot cope with the sheer number and diversity of day-to-day decision making. Decentralization of decision making is not only desirable, it is unavoidable. But decentralization *per se* solves nothing. Decentralized decisions must somehow be made compatible with central goals. And that in turn requires a system of organizational structures, performance measures, and penalties and rewards which induce decentralized decision makers to act in ways consistent with overall national plans and objectives. Participatory democracy is indeed a basic response to the overweening presence of rigid social structures and immovable bureaucracy. But participatory democracy itself will not clean up the environment, or equalize opportunity or stop the escalation of medical costs.

But how decentralized decisions can promote the social good is *par excellence* the problem with which traditional economics deals and on which it cut its teeth. Under an admittedly severe set of as-

sumptions, market prices are the duals of a set of trade-off ratios in an optimum solution. Economics, traditionally in the noncommunist countries and increasingly in the communist ones, is the discipline which analyzes how one major segment of social wants—private goods—can be satisfied by a decentralized decision making system, and which seeks to specify the characteristics which will make that system more efficient. And for the economist, the fact that modern economies often generate rapid economic growth at the expense of the environment arises not from some giant conspiracy of capitalist moguls to plunder the planet, but from the fact that the signals and incentives built into the system ignore the costs of pollution.

All of this is not to say, of course, that the private market can solve our social problems. It cannot. But it is to say that very often an analog of the market, in the form of a deliberately created structure of rewards and penalties, may be essential to handle the highly decentralized kind of decision making which has become necessary to attack the major social problems now confronting us. The concept of effluent charges in pollution control; incentive contracts and subsidies in manpower training programs; efficiency-oriented reimbursement schemes in medical insurance programs; experimental voucher schemes, individual student aid, and performance-oriented pay scales for teachers as a means of introducing incentives into stuffy and rigid educational bureaucracies; mandatory flood insurance with premiums adjusted to flood risk as a method of inducing rational floodplain development; charging the military budget for the current cost of atomic weapons and the present value of currently accruing veterans benefits as a means of inducing more rational resource allocation; providing military decision makers with incentives

to avoid "gold plating" of weapons; designing grant-in-aid programs with more flexibility and with incentives for mayors and governors to bargain with the federal government about an optimum mix of local grants;⁴ designing a system of congestion charges for airports to spread traffic, reduce delays, and provide investment signals—these are but a few illustrations of areas in which incentive considerations and market-like analogs can be used to attack frustrating bottlenecks which excessively centralized decision making places in the way of effective social programs.

These considerations lead to several others. Designing a system of institutions and incentives to channel individual decisions toward socially desirable performance implies that we have some reasonable definitions and measures of performance. The manager of the early socialist nail factory whose quota was set in terms of the *number* of nails, and who was rewarded as he made or exceeded that quota, was inevitably driven to producing *large* numbers of *small* nails, regardless of market requirements. With production quotas specified as a certain *weight* of nails, the same manager would concentrate on producing a *smaller* number of very *heavy* nails—again, regardless of market demand. In the case of medical insurance programs, incentive reimbursement schemes hold great promise for inducing more efficient use of resources and reducing the drastic escalation of medical costs. Incentive schemes can also be used to channel resources into needy areas. But concomitantly we must develop some measure of health output. We wish to reward lower costs per unit of output, not lower costs gained by undesirable reductions of

the quality of output. Small wonder that our social structures and institutions are often ineffective when all too frequently no one has devised measures of precisely what constitutes effective performance.

It is particularly in this area of defining and measuring social output that interdisciplinary cooperation and interdisciplinary research between economics and other fields can have an important payoff. Interdisciplinary research of a general nature, undertaken because interdisciplinary research—like virtue—is a good thing to pursue, is likely to combine the fuzziest elements of both disciplines. Biochemistry didn't arise because biologists and chemists in some general sense thought they should collaborate, but because the solutions of very particular problems explicitly required the kind of knowledge each could bring to bear.⁵ Similarly, in the critical task of establishing criteria for the definition and measurement of output, or performance, as part of the design of social programs, economists and other scholars have a fruitful field for interdisciplinary research on specific problems—with teaching physicians in the field of health care, with engineers and hydrologists in pollution control programs, with psychologists and educators in the case of elementary and secondary education.

Individual economists have indeed made significant contributions towards the analysis and design of institutional and incentive structures to deal with the complex problem of decentralized social decision making.⁶ But, generally, I believe, the profession has missed the boat. Together with the politicians, the public interest lobbyists, and the government bureaucrats,

⁵ See Robert M. Solow, "Science and Ideology in Economics," *The Public Interest*, No. 21, Fall 1970, pp. 101–102.

⁶ Allen Kneese's work in the field of effluent charges for pollution control is an outstanding example. See, for example, Allen Kneese and Blair T. Bower, *Managing Water Quality, Economics, Technology, and Institutions*; The Johns Hopkins Press, 1968.

⁴ Now it pays each local lobby to press for every possible "goodie," since giving up a claim to, for example, a marginal flood protection work, earns no credits towards a highly needed addition to, for example, hospital grant funds.

economists concerned with social programs have too often viewed the problem as one of specifying plans and objectives, and allocating funds. In general we have not discerned the changing nature of the problems confronting the public sector and seen the need for viewing the solutions in terms of designing institutional structures and incentive systems which automatically induce individual decision makers to move towards socially desirable ends. And yet, as I said before, this should have been the natural way for the profession to view the problem. To the extent that the *Economics* report is representative of the current concerns of economists, this is, I believe, a fair criticism.

There is another related aspect of modern social problems to which economics has much to contribute, potentially. So long as the primary concern of public policy in the economic area was the problem of overall stability and growth, we did not need detailed information about production functions in particular firms and industries. We assumed that in a reasonably competitive market system the individual firm is forced to move towards the adoption of "best practice" functions. Public decision makers did need estimates of production functions in the aggregate for devising stabilization and growth policies. But they did not need that information in the detailed and operational form which individual industrial managers require in order to produce output. But when the government itself is the producer of output, then it should act like any other producer and at least acquire knowledge of its own production function! As a larger and larger share of national resources is devoted to the production of social outputs, the provision of data and data systems from which we can deduce social production functions becomes essential. Yet we know shockingly little about social production functions. The federal government provides over \$30 billion in cate-

gorical grants to state and local governments, presumably to achieve certain specified ends. Yet we have little idea, in the case of any particular grant program, how much the grant adds to the total resources devoted to that particular purpose; how much it finances, through a substitution process, additions to spending for other purposes; and how much it relieves tax pressure on state and local taxpayers. And when the funds are spent, all too often we have even less idea of what they produce in the way of social outputs. When we push button X we don't know how much Y we get as a result, or indeed whether we get Z instead of Y.

Learning more about social production functions is a critical requirement for making social policy effective. We can impute economic growth consequences to increments of capital stock in the private economy because we have reason to believe that market pressures will force relatively efficient design and use of the new inputs by managers who have a fairly good knowledge of production functions. But for lack of any detailed knowledge of production functions in the case of social outputs, we often have little basis to choose among alternative programs or to conclude that the simple provision of additional resources will add to social output at all.

Discovering the nature of social production functions will require, in turn, major research advances.

First, knowledge about social production functions will in many cases require the careful design and subsequent evaluation of controlled experiments. The nature of the problem is often such that data generated by general surveys and administrative records will not be sufficient to discriminate between background noise and program effects.

Two kinds of experiments are possible. Most fundamentally, specific programs can be designed from the ground up to

generate experimental data, seeking to isolate input-output relationships. Politically, this may often be very difficult, since one must convince local administrators and program beneficiaries to cooperate, a far from easy task. Yet, at least in some instances, this should not be impossible. Witness the New Jersey work incentives experiment.⁷

A less ambitious controlled experiment is one which uses the natural diversity arising out of the fact that in many public programs thousands of communities use federal grant-in-aid funds in widely differing ways. The control could stem in this case from the sample imposition of carefully designed tests and measurements of inputs and outputs. Rather than accept routine administrative data (e.g., the varying educational achievement tests routinely given in schools), one might design a specific set of tests and a detailed record of inputs and outputs carried out over time in a longitudinal survey.

Paradoxically, it is precisely as economics turns its attention to such social problems as health, education, and poverty in the midst of affluence, that it must begin to approach more closely the experimental methodology of the hard sciences. In the case of the private market system, we can, as I said, assume that public authorities need not worry overly much about individual production functions. But to the extent that the public authorities themselves must choose the function, as is the case with public programs, detailed knowledge of input-output relationships

becomes critical, requiring, therefore, the kind of experimental data from which such relationships can be inferred.

This stress on knowing social production functions may seem to contradict the earlier emphasis on building incentive systems which automatically channel the interests of individual decision makers in the direction of choosing optimum input-output relationships. To the extent, for example, that a system of effluent charges is optimally levied, public authorities do not have to worry about the internal production functions of individual polluters. The more public authorities can design optimal incentive systems, the less they need concern themselves about the internal production functions of individual subordinate decision units. But this is an ideal seldom likely to be reached in most areas of social concern. As a consequence, a two-pronged approach to the problem is necessary—design of incentive systems and greater knowledge of production functions by central authorities.

In addition to learning more about social production functions there is another major requirement if economic research is to help push back the frontiers of knowledge about the major problems confronting society. We need a massive expansion in data available to researchers and research teams. The exploitation of aggregate and subaggregate data has about exhausted its potential. Almost every author in the *Economics* report points to the need for more disaggregated data in dealing with the problems in his own specialty area. Three of the six recommendations with which the report closes deal with this problem:

- creating a national data base for social and economic research, by establishing a national data system;
- establishing large scale social science research centers, one of whose functions would be collecting, systematiz-

⁷ Guy Orcutt's chapter in the *Economics* report, "Simulation, Modeling, and Data," recommends the development of large-scale, field experimentation as a major tool of research.

In the 1970 Gaither Lectures at Berkeley, Alice Rivlin explicitly considers in detail the need for and the problems associated with controlled experiments as a means of gaining knowledge about social production functions. (The lectures will shortly be published by the Brookings Institution, Washington, D.C., under the title *Systematic Thinking and Social Action*.)

ing, and making available major data systems;

- providing networks of computer facilities, with linkages to insure easy data availability to researchers at various locations.

Three different kinds of steps are necessary to provide the data base for further research efforts.

The first is the fuller availability and exploitation of existing data. Current data publication practices which make data available primarily in printed summary form is about two decades out of date. Researcher access to primary files, now possible over computer networks, would vastly increase the power of analysis, particularly by allowing simulation models to be built and operated, and by reducing the multi-collinearity problems inherent in aggregate data. Privacy problems can clearly be avoided by removal of identification and, where needed, by prohibition against outputting file data (i.e., the researcher could, through computer linkup, manipulate file data and receive back analytic relationships and summary tables, but not the raw files which were manipulated). Merging existing files—e.g., social security, internal revenue, veterans administration—on a sample basis would also multiply the usefulness of current data.

The second requirement is a large expansion in longitudinal data. The resulting combination of cross-section and time-series data on an individual basis could further multiply the explanatory power of the existing data.

The third requirement is the substantial expansion of survey research data addressed to particular problems. For reasons of continuity and economies of scale, these efforts should be centralized at a limited number of locations, but with easy accessibility of computers to researchers located elsewhere.

These points are well covered in the *Economics* report, at various places in the individual chapters and the summary recommendations. While they may indeed sound prosaic, they are essential if the discipline is to push forward its frontiers of research. In many areas those frontiers will yield only to the analysis and manipulation of large bodies of data in a sustained and continuing effort.

There are several areas in which the *Economics* report is strangely silent. It does not discuss the concept of "social indicators" (although the overall Behavioral and Social Sciences Report recommends that the government underwrite the development and continuing publication of such indicators). Without wanting to enter into the controversy about the desirability, or feasibility, of such indicators, the subject is one which has received a good bit of attention and towards which the economics profession has something to contribute. Precisely because social indicators would attempt to measure changes in "outputs" for which market prices are lacking, they could never be assembled into a single commensurable and consistent set of accounts like the GNP. Nevertheless, individual systems of accounts and summary measures, representing sub-aggregates in such fields as health, educational attainment, environmental quality, and equality of opportunity could undoubtedly be constructed. Even in these subsystems the definitions of output and the internal weights used would probably remain much more controversial and subject to much more frequent change than the concepts and internal structure of the national income accounts, because of the lack of objective, relative-value scales. Yet they might, like the national income accounts, serve a number of highly useful functions—as a framework within which to debate and formulate national objectives; as a stimulus to research seeking to

trace the effects of particular policy actions on aggregate measures of social well-being, and eventually perhaps, as a set of generally agreed-upon and politically-neutral measures of social progress.

As Professor Gurley points out in his paper, the *Economics* report is also silent on the problem of environmental quality. This is indeed unfortunate. It is unfortunate not because I believe with the ecological extremists that economic growth is a necessarily fatal cancer eating at the vitals of "Spaceship Earth," nor as Professor Gurley seems to believe, a form of robbery practiced by the rich against the limited ecological resources of the poor. Rather it is unfortunate because environmental pollution is a real problem about which economics has something of particular good sense to say. To the economist, pollution is a cost which arises from the use of the scarce assimilative capacity of the environment to absorb the waste products arising in production and consumption. Until recently, and even now to a large extent, those scarce resources were provided free, and hence overused. With the exception of some absolutely deadly residues (mercury, atomic radiation over certain cumulative amounts, etc.), the slopes of the environmental damage curves involved are not infinite; the marginal rate of substitution between environmental quality and other goods and services is not, for most citizens, vertical.

The optimum degree of environmental pollution is therefore somewhere above zero and below present levels. While pollution is a cost, its removal has a cost, and the aim of social policy should be the dual one of discovering and articulating society's (non-infinitely-sloped) trade-off between pollution and other goods on the

one hand and devising institutional procedures for balancing those preferences against pollution cleanup costs. It is quite possible that the strength of preferences for a clean environment rises more than proportionally with income—pollution control is more important to the rich than to the poor. Conversely it is also possible that the costs of pollution control will be regressively distributed (higher electricity prices, larger sewer costs, increases in automobile prices, greater proportionally for cheap than for expensive cars, etc.). As a consequence strong pollution control policies will possibly redistribute welfare against the poor. They will need to be accompanied by corrective redistributive measures if this result is to be offset.

In short, economics has much to say both about the criteria for choosing objectives in pollution control, about effective techniques to achieve control, and about the income distribution consequences of these measures. The profession can help inject some judgement and good sense into what has become—perhaps necessarily in the initial stages—a highly emotional issue. It is indeed too bad that the *Economics* report had nothing to say about the subject.

On balance then, the *Economics* report is a highly competent summary of the state of the art, and presents some well considered recommendations with respect to the extremely important data requirements which are a precondition of future progress in research. On the other hand, with some exceptions, the report is an unexciting document which fails to emphasize the very important ways in which the discipline of economics can contribute towards dealing with the urgent social problems of our time.

The State of Political Economics*

By JOHN G. GURLEY
Stanford University

This book is intended as a survey of the present state of economics. Accordingly, it contains chapters on the methods and theories of economics, including basic micro and macro theory, and on many of its applied fields, such as public finance, international trade, and agricultural economics. If one accepts the authors' own conceptions of what economics is all about, there are few grounds for serious complaint—although I can imagine a loud one from Milton Friedman to Gardner Ackley for his discussion of inflation without one—not one!—mention of the money supply. Even allowing for such grievances, the survey, for what it set out to do, is still admirable, as one would expect from such a distinguished group of theorists.

But what is and what is not surveyed is something else again. The volume appears at a time when the assumptions and methods of economics are being challenged, almost as never before, by a growing number of our younger economists, and indeed by many older ones, too. Some of the attacks have reflected dissatisfaction with the many trivial problems that economists seem to spend so much of their time on. Still others have questioned the economic goals so widely-accepted by economists, especially that of ever-increasing GNP, and a few have examined this issue within the wider framework of ecological systems. Many of the attacks on present-day economics have taken the form of radical analyses of U.S. imperialism—of how the United States profits from its leading role

in the hierarchical structure of rich and poor countries that make up the international capitalist system. Other radical analyses have examined how certain groups in the United States itself profit from the maintenance of a hierarchical class structure which produces both wealth and poverty, both privilege and oppression. These analyses, which generally conclude that capitalism is largely responsible for such social and economic disparities, accordingly deny that capitalist society is able (i.e., willing) to solve these problems. All of these studies may be said to deal with various aspects of wealth and poverty—that is with the broad social, economic, and political determinants of income and wealth distribution—both domestically and internationally. All of them strike at the very heart of conventional economics, and many of them come from young economists associated with the Union for Radical Political Economics.

The authors of this survey have disregarded all of this literature of dissent and of challenge to the basic tenets of present-day economics. As a result, an innocent reader of this survey would never know of the deep rethinking that is now going on by many members of our profession. In fact, he would gain quite the opposite impression from his reading—namely, that all is well with economics; that there is almost unanimous agreement on the fundamentals of the discipline; that economists are superbly prepared to solve what these authors call—incredible as it may seem—“newly-emerging problems” like poverty and “hot subjects” like urban decay. The dominant tone of the survey is one of great self-satisfaction and self-confidence.

* A review of Nancy D. Ruggles, editor, *Economics*, Englewood Cliffs, N.J., 1970. For helping me to improve this paper, I am indebted to Samuel Bowles and Tibor Scitovsky.

I write this review to say that such a tone is unwarranted.

I.

The conventional economics surveyed in this volume is rich in many ways, as one author after another so ably demonstrates. But conventional economics does have serious deficiencies, the most important being its distorted conception of the world we live in. This point has recently been developed by Paul Sweezy in the following way:

Orthodox economics takes the existing social system for granted, much as though it were part of the natural order of things. Within this framework it searches for harmonies of interest among individuals, groups, classes, and nations; it investigates tendencies toward equilibrium; and it assumes that change is gradual and nondisruptive. . . . [But] the world we live in is not one of harmonies of interest, tendencies to equilibrium, and gradual change. Rather, it is a world dominated by conflicts of interest, tendencies to disequilibrium, and recurring breaks in the continuity of development. A would-be science which starts with a false or irrelevant conception of reality cannot yield very significant results, no matter how refined and sophisticated its methods may be. The answers a scientist gets depend, first and foremost, not on the methods he uses, but on the questions he asks. . . . [Orthodox economics] has concerned itself with smaller and decreasingly significant questions, even judging magnitude and significance by its own standards. To compensate for this trivialization of content, it has paid increasing attention to elaborating and refining its techniques. The consequence is that today we often find a truly stupefying gap between the questions posed and the techniques employed to answer them.¹

¹ Paul M. Sweezy, "Toward a Critique of Economics," *Review of Radical Political Economics*, Spring, 1970.

It is necessary to add that the questions posed by conventional economists do, of course, include "important" and "socially relevant" problems. No one can deny that some of these economists work on poverty,

This statement describes, with a high degree of accuracy, what is and what is not in this book. The existing social system is taken for granted by every author of the survey. There is barely a mention of socialism, for example, as an alternative system. Indeed, Harberger is so oblivious to its presence that he can write of "a virtually world-wide capital market" which, in fact, excludes at least a third of the people of the world living in socialist systems. And one would never know from this book that society has a history—that industrial capitalism, its institutions, and its values have evolved from earlier societies and are now in the process of further change—for this book is about as unhistorical as it is possible to be. Furthermore, the book is long on techniques—on micro and macro techniques, econometrics, simulation, survey research, operations research, linear programming; and there is obsessive attention throughout to problems of data availability. Moreover, the book is short on social relevance, precisely because its ruling paradigm—its conception of the world—excludes power, conflict, and disruptive change within a historical setting—that is, because it excludes a large part of reality.

Political economics, as distinguished from this book's conventional economics, studies economic problems by systematically taking into account, in a historical context, the pervasiveness of ruler-subject relations in society. "The realm of the po-

pollution, racism, or whatever is considered important by the most radical members of our profession. But what is also true is that the heavy emphasis that conventional economists put on techniques tends to draw them, not to the most important social problems, which are usually terribly complex, but to those most amenable to the employment of the techniques on hand, and these are usually rather trivial problems. Furthermore, when conventional economists do address themselves to important problems, they do so through a paradigm that is such a distortion of reality that their answers are at best highly misleading ones.

litical," as political scientist Robert Tucker has written, "[is] the realm of power and authority relations among people. Such relations between man as ruler and man as subject occur not alone through the medium of the state—they occur in virtually every other form of society as well, starting with the family. All established human groups and institutions," Tucker goes on to say, "have their inner structure of authority, their pattern of ruler-subject relations. If they stand outside the institutional structure of the state, they are not on that account nongovernmental. Rather, they belong to the sphere of private, as distinguished from public government. Government itself, in the elementary sense of rulership of man over man, is pervasive in human society."²

I will add that it is these pervasive relations of domination and servitude, these relations of power and authority that lead to conflict, disharmony, and disruptive change. A political economist sees these power structures and puts them at the forefront of his analysis; a conventional economist—who sees only a society of free, self-interested economic men interacting as equals in the marketplace—does not.

To extend the discussion of these differences, the conventional economist, because he sees harmonies of interest almost everywhere, can visualize himself as a neutral technician, applying his techniques as objectively as possible. If class or group conflicts do not exist, one can, of course, work for the "general interest." The political economist turned radical, on the other hand, not only studies economic problems within the historical context of ruler-subject relations, but he actively takes the side of the poor and the powerless, and he generally sees the system of

capitalism as their oppressor.³ He believes, therefore, that the conventional economist not only fails to take account of relations of power and authority, and so fails to grasp the most socially relevant aspects of the problem, but, by being so blind to class interests and so caught up in his data and his techniques, he in effect supports a system that maltreats large numbers of people.⁴

II.

These observations apply with full force, I believe, to Hollis Chenery's chap-

³ One can be a political economist without being a radical political economist, though there are probably few economists who fall into such a category—say, of right-wing political economist. These two types of political economist, though opposed to each other on political grounds, nevertheless have much in common in that they share essentially the same paradigm. The relation of either type of political economist to a conventional economist, whether the latter is of conservative or liberal persuasion, is necessarily less close because their paradigms are so very different.

⁴ Consider a concrete problem. In looking at the tradeoff between inflation and unemployment, a conventional economist would consider this to be a technical problem of quantifying a functional relationship, which might be solved with good data and appropriate econometric techniques. He would, therefore, spend most of his time in trying to get the data into the best form for rigorous testing of some hypothesis about this tradeoff, and in the interpretation of his results.

A radical political economist, looking at the same problem, would also be interested in its quantitative and theoretical aspects, but in addition he would ask why it is that the institutional arrangements of advanced capitalism require that we have either inflation or unemployment. He would further ask who gains and who loses by the various tradeoffs between inflation and unemployment, and he would assume conflicts of interest among groups in the society and a highly unequal distribution of power among these groups to influence decision making in this area. These thoughts might lead him to investigate whether the Nixon Administration's strong emphasis on stopping the inflation, at the cost of more unemployment, might not serve a powerful group of bankers and industrialists whose financial interests heavily depend on the continuing dominance of the United States in the world capitalist system and hence on the enduring strength of the dollar. He would, of course, at the same time reflect that unemployed blacks and chicanos have very little power over such decisions compared to, say, David Rockefeller.

² Robert C. Tucker, *The Marxian Revolutionary Idea* (W. W. Norton & Company, 1969), p. 80.

ter on economic development. In these pages, there are no power relations of domination and servitude, no basic conflicts between countries or within countries, no social disharmonies—only some technical difficulties. The heavy emphasis, therefore, is on research techniques to overcome the technical difficulties—on input-output analysis, cost-benefit studies, linear programming models with shadow prices—and on the availability of data to feed these techniques. In Chenery's world, and Harberger's too, international society consists of free, self-interested nations, some rich and some poor, each with a more or less socially homogeneous population, interacting as equals on international markets.

Now Chenery's paradigm omits enough of reality to make it likely that he would overlook some important questions about economic development. What his model permits him to see are narrow economic barriers to the development of poor countries—inadequate savings, inappropriate relative prices, shortages of entrepreneurs, imbalances between this and that, and so on. These are, without question, important. But what he cannot see are factors of at least equal importance—the social, political, and broader economic barriers to development that arise out of class conflicts of interest. Thus, he does not discuss whether economic development is even possible, given the social and political requirements of the domestic oligarchies and the financial interests of the foreign investors (indeed, foreign investors are not discussed at all). Who gains from economic development, and who loses? And which groups have the power to make the decisions? These are questions of political economics, questions of power, class conflicts, and resistances to change. Chenery does not ask them because he sees only harmonies of interest; everyone is working for development, the rich and the

poor together, and everyone stands to gain.

But this is often not true. A thoroughgoing development program that reaches deeply into all the stagnant backwaters of society is clearly in the interests of the poor in underdeveloped countries, but it is not necessarily the top priority of those who rule them, nor of the foreign corporations engaged in exploiting their resources. The over-riding goal of the ruling oligarchies is usually to maintain, and if possible to enhance, their privileged positions of wealth and power in the economic, political, and social life of the country. The principal requirement of the foreign corporations is for a stable and highly favorable environment for their investment and trading activities.

A broadly-based and deeply-penetrating development effort is a serious threat to both of these priorities, for such an effort often demands nothing less than a social revolution. Landlords may lose power and with it the ability not only to exploit peasants economically but to oppress them politically and socially. Peasants and workers may have to be educated, awakened, and organized to fulfill the tasks of a serious development effort, and this awakening may prove subversive in the end to the ruling class's privileged positions. In many of these countries there is already organized armed resistance by the poor to their governments, which could be greatly strengthened by the newly-released energy of a general shake-up of the society. Consequently, it is not surprising that there are often strong alliances between the domestic oligarchies and the foreign corporations for the promotion of stable environments, for this serves the priorities of both groups. If economic development is encouraged at all, it is advanced within the safe framework of the existing power structure, and this calls for moderate policies of marginal improvements in agricul-

ture and light industry, roadbuilding, family planning agencies, and the like—policies that will not rock the boat. Such development efforts are more likely to serve the financial interests of the wealthy than the basic needs of the poor.

Furthermore, it is not correct to assume that a major goal of successive U.S. governments has been the economic development of poor countries. It has not. The principal drive has been, as Harry Magdoff has put it, "to keep as much as possible of the world open for trade and investment by the giant multinational corporations."⁵ And this aim is often vigorously supported by the international financial institutions. Sometimes the pursuit of this goal requires economic aid from the United States to poor countries; but sometimes it does not. Economic aid is most likely to be forthcoming when it serves the above goal by providing much needed overhead capital for the foreign investors and by helping to quiet general unrest and to defuse potential insurrections. However, economic aid is not always necessary to quell unrest, for potential threats to the ruling classes can also be met by military aid to them, counter-insurgency operations against dissident forces, and, if worst comes to worst, military intervention itself—which has been used by the United States around 40 times since the 1890's. Militarism in the United States is integrally related to U.S. imperialism.⁶ These are alternative means to a given end, and economic aid, as only one of the means, may or may not play a large role, depending on the circumstances in each case. There is thus a powerful alliance within the international capitalist system working for ends that are often op-

posed to the basic needs of the poor.⁷

Since Chenery fails to discuss these obstructions to the welfare of the poor, his chapter excludes a rapidly growing body of literature which, in fact, addresses itself to just such problems and which contains some of the most promising work now being done in the development field. Among others, I have in mind A. G. Frank's studies of Latin American development; Michael Tanzer's book on the political economy of international oil; Harry Magdoff's empirical researches on U.S. imperialism; Paul Baran's path-breaking historical analysis of world economic development; Paul Sweezy's many contributions to the same topic; Robert Heilbroner's and Gunnar Myrdal's works, which delve into the social and political aspects of development; and the expanding literature on imperialism and development contributed by some of the younger economists.⁸ In a book which purports to

⁷ Conventional economists, of course, deny this, because they believe that the rich are helping the poor. Radical economists, on the other hand, find enough evidence to assume that the rich are helping themselves; in the process the poor may be helped to some extent, or they may not.

⁸ I shall list only a few of these writings. A. G. Frank, *Capitalism and Underdevelopment in Latin America* (Monthly Review Press, 1967); Michael Tanzer, *The Political Economy of International Oil and the Underdeveloped Countries* (Beacon Press, 1969); Harry Magdoff, *op. cit.*; Paul Baran, *The Political Economy of Growth* (Monthly Review Press, 1957); Paul Sweezy, *The Theory of Capitalist Development* (Oxford University Press, 1942); Robert Heilbroner, *The Great Ascent* (Harper Torchbooks, 1963); Paul Baran and Paul Sweezy, *Monopoly Capital* (Monthly Review Press, 1966); Edward Boorstein, *The Economic Transformation of Cuba* (Monthly Review Press, 1968); Eric Williams, *Capitalism and Slavery* (University of North Carolina Press, 1944); James O'Connor, "International Corporations and Economic Development," *Science and Society*, Fall 1969; Dudley Seers, "Big Companies and Small Countries," *Kyklos*, XVI, No. 4, 1963; Thomas Balogh, "The Mechanism of Neo-Imperialism," *Oxford University Institute of Statistics Bulletin*, August 1962; Tom Weisskopf, "Capitalism, Underdevelopment, and the Future of the Poor Countries," Unpublished, 1970; Richard D. Wolff, "Modern Imperialism: The View from the Metropolis," *AER*, May 1970; Theotonio

⁵ Harry Magdoff, *The Age of Imperialism* (Monthly Review Press, 1969), p. 14.

⁶ Harry Magdoff, "Militarism and Imperialism," *American Economic Review*, Papers and Proceedings, May 1970.

survey the state of economics, any chapter that slights such a large and challenging body of work cannot be considered entirely representative of what economists are doing, nor can it be considered to have posed the most significant questions about its subject.

III.

As I noted previously, radical economists have also been interested in the political economics of wealth and poverty right here in the United States—and that means, to repeat, that they have analyzed their topics within a historical context of ruler-subject relations. These topics include how the capitalist system produces simultaneously both wealth and poverty; the economic, political, and social power bases of the rich; the political economics of poverty, of slavery, and of black ghettos; and, in general, the broad social determinants of income and wealth distributions in our contemporary society. It goes without saying that conventional economists too, over the years, have contributed much to our knowledge of income and wealth disparities.⁹

Dos Santos, "The Structure of Dependence," *AER*, May 1970. I have attempted to summarize some of the ideas in these studies in "Economic Development and Economic Imperialism," *Plain Wrapper*, April-May 1969, and in "Economic Conversion and Beyond," *Industrial Management Review*, Spring 1970.

⁹ The following is a selection of these writings, mostly on the radical side. Eugene Genovese, *The Political Economy of Slavery* (Vintage Books, 1965); Gabriel Kolko, *Wealth and Power in America* (Praeger, 1962); William K. Tabb, *The Political Economy of the Black Ghetto* (Norton, 1970); James C. Knowles, *The Rockefeller Financial Group* (mimeo., 1970); F. Lundberg, *The Rich and the Super-Rich* (Lyle Stuart, Inc., 1968); Victor Perlo, *The Empire of High Finance* (International Publishers, 1957); Robert J. Lampman, *The Share of Top Wealth-Holders in National Wealth, 1922-1956* (Princeton University Press, 1962); Charles E. Silberman, *Crisis in Black and White* (Vintage, 1964); Paul Sweezy, "Interest Group in the American Economy," *The Present as History* (Monthly Review Press, 1953); J. G. Gurley, "Federal Tax Policy," *National Tax Journal*, September 1967; Daniel Fus-

But this survey is strangely silent on such matters. Not only does it disregard radical economists' work in this area, but it does not deal in any serious way with the plain old topic of income distribution. In fact, the topic is so chopped up in the volume that it is impossible for any author, even if he wanted to, to come to grips with the heart of the issue. Caves considers why some corporations get higher profit rates than others. Ruttan wonders briefly about income distribution in agriculture, and Rothenberg refers to inequities in urban areas. Reder asks why some workers earn more wages than others. After all this, there is nobody left to ask questions about the overall distribution of income; and, of course, the assumption that capitalists deserve the profits they get and the wealth they have is never challenged at any point in this survey. Solow, for example, can write about the competitive model in this way—"There is a given set of households owning specified assets (including the ability to supply certain types of labor and *the right* to a share of the profits of each firm) . . ." ¹⁰—presumably without batting an eye, for in that formulation everyone sounds pretty much equal. Solow takes

feld, "The Basic Economics of the Urban and Racial Crisis," *Conference Papers of the Union for Radical Political Economics*, December 1968; Stephan Michelson, "On Income Differentials by Race: An Analysis and a Suggestion," *ibid.*; Barry Bluestone, "Black Capitalism: The Path to Black Liberation?", *Review of Radical Political Economics*, May 1969; Raymond Franklin, "The Political Economy of Black Power," *Social Problems*, Vol. 16, No. 3; Victor Perlo, "People's Capitalism and Stock Ownership," *American Economic Review*, June 1958; Margaret Benston, "The Political Economy of Women's Liberation," *Monthly Review*, September 1969; Donald Light, "Income Distribution: The First Stage in the Consideration of Poverty," *Occasional Papers of the Union for Radical Political Economics*, December 1969; Stephan Michelson, "The Economics of Real Income Distribution," *The Review of Radical Political Economics*, Spring 1970.

¹⁰ Robert Solow, "Microeconomic Theory," p. 4 in mimeographed version. My emphasis.

the existing economic system so much for granted that it limits his imagination—"... one would guess that the elimination of substantial unemployment—let alone mass unemployment—does as many good things to the income distribution as any direct measure one can imagine."¹¹ Well, any radical economist can imagine a direct measure that would do even better things—expropriation of the capitalist class and the turning over of ownership of capital goods and land to all the people. That, of course, sounds wild—unimaginable—to anyone who does not question the existing social system.

But if one is concerned with gross inequalities of income and wealth, as radical economists are, it is best to begin with property incomes, and especially with the profits going to private owners of capital goods. Are these profits deserved? Radical economists analyze this problem, as they do most problems, from the standpoint of power and conflict within a historical setting. To begin with, as Joan Robinson has pointed out, capital goods *are* productive but that does not justify the assignment of their productivity to private owners. There is nothing productive about land and machinery being privately owned, and it is certainly no sacrifice or burden to own such wealth. In fact, the present ownership of capital has a historical background that is scandalous, as our Indian, chicano, and black contemporaries have been trying to explain to us in recent years.

Furthermore, the sacrifice associated with current saving by the wealthy is more truly a burden of the poor whose very poverty—whose forced abstinence from consumption, to put it politely—allows such wealth, from which most of the saving comes, to continue to exist. It is not the abstinence and waiting of the rich

that should be rewarded but rather the general deprivation of the poor. Also, the long established economic argument against private acquisition of returns on land has not been forgotten by radical economists, even though its political implications are seldom discussed any more in mainstream economics. In short, radical economists reject the attempts by conventional economists "to raise profits [and other property incomes] to the same level of respectability as wages."¹²

As radical economists see it, the shares of national income going to workers and to property owners are largely determined by the relative power of the two groups, although relative factor supplies set limits within which the power exerts itself. But that is hardly the whole story, for even if, in some sense, workers got something close to their marginal products, the marginal products themselves are considerably influenced by the very fabric of the society in which people work.

Adam Smith was acutely aware of this when he suggested that labor specialization, while increasing productivity in some narrow sense, is often at the expense of the worker's general intelligence and understanding. It was his view that "the man whose whole life is spent in performing a few simple operations . . . generally becomes as stupid and ignorant as it is possible for a human creature to become." The difference between the most dissimilar of human beings, according to Smith, is not so much the cause of the division of labor as it is the effect of it. Consequently, while an economy might gain from the division of labor in some small sense, it could lose in the larger sense of creating

¹² Joan Robinson, *Economic Philosophy* (Anchor Books, 1964). See also, by the same author, *An Essay on Marxian Economics* (St. Martin's, 2nd ed., 1966); *Economics: An Awkward Corner* (Pantheon, 1967); and *The Accumulation of Capital* (St. Martin's, 1965).

¹¹ *Ibid.*, p. 15 in mimeographed version.

men who are little more than passive and unreasoning robots.¹³

In general, the marginal products of workers are not independent of the opportunities the society affords to all its citizens to become fully participating members of that society—to become fully-awakened and myth-free individuals. In particular, in our own society, these marginal products are not independent of the inferior status accorded to women, the racism that is so widespread, the class-oriented education system, or the officially-sponsored efforts to divert the energies of millions of people into imaginary worlds.¹⁴

¹³ Adam Smith, *The Wealth of Nations* (Irwin, 1963), Vol. II, pp. 284–85. I have pursued these and associated ideas in "Maoist Economic Development," *The Center Magazine*, May 1970. See also Nathan Rosenberg, "Adam Smith on the Division of Labour: Two Views or One?" *Economica*, May 1965.

¹⁴ This general theme has been developed by several radical economists. For example, as Eugene Genovese has shown, in the ante-bellum South, the maintenance of slavery as an entire social system required the slaveowners to keep their slaves stupid. This in turn meant that the level of technology had to be kept down to the low quality of the labor force. Thus, slaves used the cruder "nigger hoe" instead of the better but more easily breakable "yankee hoe," and they were forced to work with mules and oxen instead of faster horses because the former "could more easily withstand rough and perhaps vengeful handling." Further, the more complicated tools and machinery could not be used because the social system militated against encouraging slaves as specialists to service the equipment. Thus, the entire social system required stupid slaves, and stupid slaves required crude tools. Educated, technically-trained, and wide-awake slaves, although they would have been far more productive in the fields, would have threatened the slave society as a whole. So slaves got low pay, it could be said, because their marginal product was low, but it was low because it was imposed on them by the social requirements of the ruling class. See Eugene D. Genovese, *The Political Economy of Slavery* (Vintage Books, 1965), especially pp. 3–10, 43–69, 244.

If today's conventional labor economist had lived in that society, judging from Reder's survey of the field, he would not have discussed any of these fundamental forces determining the income-share of slaves—including the basic fact that slaves did not own the means of production—but would have concentrated instead on slave X receiving more than slave Y because he invested capital in himself and learned the alphabet; or slave X receiving more because slave-

These are all subjects which are not only worthy of discussion by economists, but which have, in fact, been discussed, and never more than in recent years, but none of them finds a place in this survey.¹⁵

IV.

What does find a place in every nook and cranny of the survey is the assumption that the source of human welfare is commodities, i.e., GNP. In these terms, greater welfare requires more "productive" workers—workers molded by capitalist institutions to the narrow pursuit of producing ever more to consume ever more. Nowhere in this survey does any author discuss the implications of these limited goals of capitalism for the human development of the people involved.¹⁶

Instead, we learn from the survey that *the* economic goal is growth—and that's that. It is really unbelievable that no author questions this. For years there have been discussions on the high costs of growth¹⁷ (hasn't Boulding proposed to

owners had a "preference" or "taste" for his lighter colored skin. And if a Caves had been on the scene in those days, he would have carefully examined whether slaveowners were trying to maximize profits, why slaveowner A got a higher profit rate than slaveowner B, and the implications of this for differences in market conduct and performance between the two slaveowners. And Morgan presumably would have conducted consumer surveys to see what makes slaves tick. If the conventional economists ran true to form, the emphasis of these analyses would have been on harmony of interests and possibilities for reforming the worst abuses; the analyses would not have penetrated into the heart of the slave society itself.

¹⁵ However, as I have already explained, this is not unexpected because the paradigm of conventional economists, which takes the existing social system for granted, greatly reduces their chances of seeing power structures, social conflict, and strong resistances to change—all of which form the very core of the problems of wealth and poverty, as radicals see them.

¹⁶ See Herbert Gintis, "Neo-Classical Welfare Economics and Individual Development," *Occasional Papers of the URPE*, July 1970.

¹⁷ I will note only two of the many studies that question our singleminded dedication to high growth rates: E. J. Mishan, *The Costs of Economic Growth*

change GNP to Gross National Cost, and others to Gross National Pollution?) and even a great deal of speculation about whether growth, even if wanted, is, in fact, possible for very much longer. This topic, which is a serious challenge to both conventional and radical economists, deserves some place in a survey of what's going on in economics and where economics is going.

In its most comprehensive form the argument is based on the burgeoning growth of population relative to resources and food; on the limited waste-absorbing capacity of the soil, water, and atmosphere; and on the special environment-destroying potential of newly-developed technologies, like nuclear radiation.¹⁸ Among economists, the argument in this form was stated most cogently several years ago by Kenneth Boulding, but it has recently been put into even more dramatic form by Robert Heilbroner, in this way:

[In] our Spaceship Earth . . . sustained life requires that a meticulous balance be maintained between the life-support capability of the vehicle and the demands made by the inhabitants of the craft. Until quite recently, those demands have been well within the capability of the ship, both in terms of its ability to supply the physical and chemical requirements for continued existence and to absorb the waste products of the voyagers. . . . It is only in our time that we are reaching the ceiling of earthly carrying capacity, not on a local but on a global basis. Indeed . . . we are well past that capacity, provided that the level of resource intake and waste output represented by the average American or European is taken as a standard to be achieved by all humanity. To put it bluntly, if we take as the price of a first-class ticket the resource requirements of those passengers who travel in the Northern Hemisphere

of the Spaceship, we have now reached a point at which the steerage is condemned to live forever—or at least within the horizon of the technology presently visible—at a second-class level; or at which a considerable change in living habits must be imposed on first class if the ship is ever to be converted to a one-class cruise.¹⁹

This analysis has many implications for both the rich and the poor, but I have space to mention only one of them. In attempting to achieve economic development, today's poor around the world have a few advantages over their predecessors of a few hundred years ago—the advantages of increased knowledge and of better technologies. But they are at a disadvantage in several crucial ways. First, they do not have the opportunities which the early Europeans and Americans had of gaining huge flows of economic resources through plunder, slavery, and fraud. Second, I have argued that they are burdened by a powerful alliance within the world capitalist system which works for ends that are often opposed to their basic needs. And, finally, to the extent that the above "ecological argument" is correct, the present poor are further condemned by the fact that the rich states are hell-bent on a growth path of demanding and producing ever more GNP, and in the process using up exhaustible resources that the poor must have if they are ever

¹⁸ Heilbroner, *op. cit.* This article also appears in *Between Capitalism and Socialism* (Vintage, 1970). It was a review of a book by Paul and Anne Ehrlich, *Population, Resources, Environment* (W. H. Freeman & Co., 1970). The above passage of Heilbroner is a paraphrase of the Ehrlichs' argument.

See also Kenneth Boulding, "The Economics of the Coming Spaceship Earth," in Henry Jarrett (ed.), *Environmental Quality in a Growing Economy* (Johns Hopkins Press, 1966). Boulding's article may also be found in Garrett de Bell (ed.), *The Environmental Handbook* (Ballantine Books, 1970). For much the same argument, see Herman E. Daly, "The Canary Has Fallen Silent," *N.Y. Times*, October 14, 1970, p. 43.

(Praeger, 1967) and J. K. Galbraith, *The New Industrial State* (Houghton-Mifflin, 1967).

¹⁹ This formulation is based on that of Robert Heilbroner, "Ecological Armageddon," *The New York Review of Books*, April 23, 1970.

going to raise their own living standards to halfway decent levels.

So the poor of the world have problems on their hands, and one of them is the rich and their pecuniary ways. The rich might reform. Or the poor, in their hundreds-of-millions, might one of these days become an organized revolutionary force against the rich—and so, incidentally, disprove my first point that they have no opportunity these days for plunder.

V.

At the beginning of this paper, I argued that conventional economists unduly emphasize, at the expense of more important things, methods and techniques and the data to feed them. I repeat this because I want it as immediate background to what I now wish to read to you: the final six recommendations of this panel of economists who have surveyed the entire field of economics. Here they are:

1. The panel recommends that the federal government direct immediate attention to the development of a data base for economics and social research by the establishment of a National Data System . . .
2. The panel recommends that private foundations and government agencies aid in the establishment and support of large scale social science research centers which can provide continuing facilities for economic and social research . . .
3. The panel recommends that universities and research centers cooperatively with the federal government assume the responsibility for providing computer facilities for economic research . . .
4. The panel recommends that private foundations and government agencies together with universities and research centers provide for the funding and organi-

zation of long-term economic research efforts . . .

5. The panel recommends that departments of economics reorient graduate training towards providing more competence in research methods and more research experience . . .

6. The panel recommends that professional societies and research organizations accept greater responsibility for and give explicit consideration to improving the channels of communication in the profession . . .²⁰

I would like to counter with a single recommendation of my own—that the economics profession, instead of worrying so much about national data systems, research money from foundations and the government, computer facilities, and research techniques, pay more attention to the economic implications of the fact that the United States is heavily involved as the dominant power in a hierarchical international capitalist system. Once this fact is recognized, economists are bound to have interesting and important things to say about U.S. objectives within this system, the various means employed by successive U.S. governments to achieve these objectives, the sources of power in the U.S. which promote such ends, the impact of all this on militarism and on stabilization policies in this country, and the economic implications for poor countries of being subordinate members of such a system. One doesn't need fancier computer facilities or national data systems to make important contributions in these areas, and one doesn't need more research money to gain a better understanding of the awful problems that the poor and oppressed people of the world are up against.

²⁰ Chapter 17, in mimeographed version.

DISCUSSION

ROBERT M. SOLOW: I have no substantial disagreement with Charles Schultze. He is, perhaps, what you might call a little justifiably unfair to the *Report*. It was, after all, supposed to describe the present state of economics—with outsiders as the main audience—not to blaze new paths. I'm afraid it is an unexciting document. (It was written by a Committee!) It suggests no double helix or miracle rice, probably because none of us could see such a thing on the horizon in economics. (Those are Famous Last Words, I realize.)

Some of the things Schultze talks about were indeed mentioned in the *Report*. I had a paragraph in my chapter on "Microeconomic Theory" suggesting that the study of commodity taxes and user charges to accomplish allocational purposes would be an important research topic for the future. I mentioned urban economics as a major source of such problems. I had traffic congestion rather than pollution in mind. I might as well confess that I discovered the interesting and important work of Kneese and his associates only after I had written my chapter; otherwise I would certainly have mentioned it prominently. It carries an important message that economists have to transmit to others.

Schultze may be a little less than fair when he says that economists have tended to regard the design of public programs as a problem of specifying objectives and regulations rather than as one of providing correct incentive schemes for decentralized decisions. There may be some truth in that—he is in a better position to know than I am—but it is at best a partial truth about a profession that has invented and spread the whole idea of shadow-prices in the last 20 years.

Generally, I see the immediately promising directions of theoretical and applied research where Schultze sees them. Despite the flurry over monetarism-and-fiscalism and natural-rate-of-unemployment vs. permanent-trade-off, I think these are essentially minor issues in which good sense will eventually win out. As Schultze argues, microeconomics is due for a revival as problems of regulation and decen-

tralized management of the economy come to the fore. The Averch-Johnson literature is another example of the analysis of the subtle distortions caused when the regulatory process provides improper incentives.

For some reason, economists often meet radical blasts like Professor Gurley's with a kind of embarrassed silence. Perhaps that is because one tends to swallow the notion that doing conventional economic analysis somehow identifies you with the System and its inequities, not to mention its foreign policy. I think that notion is false. If you want to work for reform of the System, or even for revolution, that is no excuse for abandoning the usual standards of scientific work. To do so is to add ignorance to injury. I have known Professor Gurley for a long time. I have always taken him seriously and I propose to do so now. But I'm going to take a tough line, just as he did.

The *Report* under review is a survey of the present state of economics, and clearly not an exhaustive survey. Gurley's review is much more an attack on the present state of economics than an attack on the *Report*. The *Report* itself is faulted only if the body of radical political economics is so large and so important that it demands mention. But I don't believe that is the case. Radical economics may conceivably be the wave of the future, but I do not think it is the wave of the present. In fact, to face the issue head on, I think that radical economics as it is practiced contains more cant, not less cant; more role-playing, not less role-playing; less facing of the facts, not more facing of the facts, than conventional economics. In short, we neglected radical economics because it is negligible.

First I want to say that I think radical economists have corrupted Thomas Kuhn's notion of a scientific paradigm, which they treat as a mere license for loose thinking. If you look at Kuhn's examples—all from natural science, of course—you will see that they represent well developed models or frameworks for thought. Some of his examples are Newtonian dynamics, Copernican astronomy, Ptolemaic astronomy for that matter, wave

optics, etc. In this sense, neoclassical economics is pretty clearly a scientific paradigm. It may be a bad one, or a worn-out one, or it may have served to advance the interests of the capitalist class, but it is the sort of thing Kuhn means. As far as I can see, radical political economics is no such thing. It is more a matter of posture and rhetoric than of scientific framework at all. (Classical Marxian economics is a different matter, of course, though there is a problem about its relation to the classical paradigm. Anyhow that is another story. "It is no accident," as they say, that modern radicals are not much interested in the old man who wrote *Capital*. They are much more interested in the Hegelian Marx, and the author of the Paris Manuscripts of 1844.) Maybe I can put my point in Kuhn's language. The function of a scientific paradigm is to provide a framework for "normal science." But there is little evidence that radical political economics is capable of generating a line of normal science, or even that it wants to.

Here are some examples of what I mean. Professor Gurley says: "As radical economists see it, the shares of national income going to workers and to property owners are largely determined by the relative power of the two groups, although relative factor supplies set limits within which the power exerts itself." Am I to presume from this that there are studies of time series that show that short-run fluctuations in distributive shares reflect short-run fluctuations in the distribution of power in society? This would mean that workers are more powerful when there is a lot of unemployment than they are when there is very little, because the share of wages is highest when the economy is most depressed. Or has it been found in many countries that the direction of long-run change in distributive shares corresponds to the long-run trend in the independently measured distribution of power in society? Or are there perhaps cross-section studies among industries showing that the share of wages in value added is highest in industries where the power of the workers is highest, and lowest where the social power of capitalists is highest? Or is it demonstrable that international differences in functional income distribution correspond more or less

closely to international differences in the relative power of workers and property owners?

According to Gurley, contemporary underdeveloped countries "are at a disadvantage in several crucial ways. First, they do not have the opportunities which the early Europeans and Americans had of gaining huge flows of economic resources through plunder, slavery and fraud." I presume this means that it is well understood why Spain and Portugal, which must have had more than a fair share of plunder, slavery, and fraud, are today little more than underdeveloped countries themselves, while Sweden and Australia (and, because they were late on the scene, perhaps also Germany and France), which did not, are well-developed high-income economies. I presume that there are estimates of the extent to which Japan's success comes from its earlier adventures in China and how much from other sources. I presume that radical economists have a way of finding out how much England profited from India, from central Africa, and from the Atlantic colonies. And I hope it is understood that I can doubt that answers to these questions exist without in any way favoring plunder, slavery or fraud.

Professor Gurley says, meaningly, that a radical economist might "investigate whether the Nixon Administration's strong emphasis on stopping the inflation at the cost of unemployment might not serve a powerful group of industrialists whose financial interests heavily depend on the continuing dominance of the U.S. in the world capitalist system." Apart from his current pressing need to explain why the Nixon Administration appears to have changed its mind early in November, the radical economist has presumably calculated whether the strength of the dollar is worth enough to those bankers and industrialists to offset the large drop in profits they experience whenever there is strong emphasis on stopping inflation at the cost of unemployment.

As you all know, none of these calculations has been done by radical economists. And when and if any of them is done, you know who will do it—some poor damned graduate student in some conventional department, supervised by some conventional professor of conventional economics. That's why I don't

think a survey of the current state of economics needs to pay a lot of attention to radical political economics.

May I say a word in my own defense? Professor Gurley charges me with having, without batting an eye, described a model in which each household has a given capacity to supply specified kinds of labor, and a right to a specified share of the profits of each firm. He says this in a tone of hushed shock, for all the world as if he were a Victorian literary critic announcing that in Chapter 12 of my novel I actually say, without batting an eye, that Cedric put his hand on Gwendolen's breast! In the first place, I was describing the competitive model, which does just as I said. Secondly, the competitive model was intended as a description of capitalism. Third, the right in question is patently a legal right, not a moral right. And fourth, with trivial modification, the same model could be adapted to describe many alternative social-economic systems because the surplus generated in firms has to be distributed somehow.

Moreover, when I remarked that ending unemployment did as much good to the income distribution as any imaginable direct measure, I was plainly referring to any direct measure that one could imagine actually happening between 1960 and 1965. Talking about expropriating the capitalist class in that context is what I mean by role-playing. By the way, does Gurley think that the phrase "turning over of ownership of capital goods and land to all the people" says anything precise? Does he think the distributional consequences are implicit in the phrase? Does he have any clear idea what the income-distributional consequences would be, or have been in the various socialist countries who have done this in a wide variety of institutional ways? And why, in this connection, does he stop talking about "power and servitude" and ruler-subject relations, which I thought were more or less ubiquitous, and are certainly not noticeably absent in places where the capitalists have been expropriated?

Finally, I must say I don't think it's wrong for a survey of economics to talk mostly about technique and the availability of data. There is a difference between a survey of medicine

and a survey of health, or between a survey of meteorology and a survey of the weather. So why not between a survey of economics and a survey of economic life? I presume that knowledge of technique and acquaintance with data *are* what distinguish an economist from somebody else. I'd even have thought that it is the command over certain techniques and the familiarity with certain bodies of data that distinguish the radical economist from the radical journalist, or the radical sociologist, or the radical clarinetist. What else?

ROBERT L. HEILBRONER: As always, there is too much to say and too little time to say it in. I shall therefore proceed in telegraph style, with little or no attempt to connect my comments into a smoothly knit whole.

1. Both commentators agree that the BASS report is a bland document. Curiously, Gurley, whose criticism of the report as a whole is much more radical than Schultze's, finds the contents of the volume, within their own context, "admirable." I suspect this is because he is not really interested in them. Schultze, who accepts the socioeconomic premises on which the report stands, has far more penetrating criticisms to offer in terms of the sins of omission for which such a survey on conventional economics can be properly taxed. I was especially struck with his concept of social (public agency) production functions as a field in which knowledge is virtually absent and sorely needed.

2. Schultze is eager to use economics as a technical discipline for the solution of important policy issues and he is confident that the discipline can produce important results. The question to which he does not address himself, however, is how we are to delimit these issues of policy. He accepts a definition of economics in which "current social structures" are parametric, although he speaks about modifying "social behavior and social institutions." This involves him in a dilemma. What are the possibilities of behavioral and institutional change within a given social structure? The question remains unasked. As a result, the problem that haunts Gurley remains unanswered. How does *power* enter the picture? How do we reconcile the dual purposes of ev-

ery economic system, to wit, its *social* function as an instrument for the provision of goods and services, and its *political* function as the distributor of differential rewards to different social classes? It is in its uneasiness before this problem—indeed, in its unwillingness to confront the problem—that conventional economics meets a fundamental barrier in terms of fulfilling the very tasks that Schultze demands of it.

3. Further by way of comment on Schultze's paper, I am increasingly rendered uneasy by definitions of economics that contain the idea of maximizing economic output. There are two reasons for this. First, maximization is a behavioral prescription that is exceedingly difficult (perhaps impossible) to define objectively in conditions of less than perfect competition and over periods longer than very short run. Second, we have become more and more aware that externalities constitute a major source of economic and noneconomic trouble, so that policies that seek economic maximization alone may constitute disasters for society as a whole.

4. Gurley's dissatisfaction with the BASS report lies on entirely different grounds. If I may paraphrase what I believe to be his complaint, it is that conventional economics accepts as parameters the main institutional forms of our social system and directs its energies toward explaining the interaction of the variables within that system. Radical economics wants the discipline to turn its analytical abilities on the parameters themselves and to explain *them*, not just their effects.

I find myself in much sympathy with this general desire to widen and deepen the range of what is called economic analysis. Recently I had occasion to visit the economics department of another university whose chairman spoke of the importance of teaching the "lasting substance" of economic analysis rather than the "ephemeral data" of social reality. I asked him which he thought was more enduring—the analytical content of economic theory over the last 25 (not to say 50!) years, or the ephemeral data of income distribution, industrial concentration, or world development differentials. (He said, hm-m-m-m, he thought I had a point there.)

5. Over-aggregation is the *bête noire* of the conventional economist, and quite correctly. But I would suggest that there is an analog in radical or political economics, which I would call over-generalization. Is it capitalism alone—as the radical economists seem to maintain—that is responsible for the ills of racism, alienation, exploitation, bureaucratic indifference, etc.? If so, then we should expect to find these problems present in every country in which there exists a concentrated private ownership of property and a major reliance on the market system, and absent in those countries in which the means of production are publicly owned and in which planning has asserted its preeminence over the market. I then ask why life in Sweden, Denmark, England, and Norway is so different from life in the United States, not to mention Japan or the Union of South Africa; and I ask as well what we are to say about anti-Semitism in Poland, the suppression of intellectual dissent in Cuba, the invasion of Czechoslovakia by the U.S.S.R., etc. In a word, unless we apply to the study of parameters the same exacting criteria that we apply to the study of variables, the result of radical economics will be wishful thinking, not social science.

7. There is a crying need to penetrate the catch-words of conventional economics that spread a scientific-sounding gloss over ignorance. I nominate "opportunity cost" to head the list. There is no less of a need to apply an equally penetrating examination to the words of political economics that spread a humanistic-sounding gloss over areas of equal ignorance. I nominate "the people" to head that list.

8. Finally, the question of where we go from here. Within the framework of conventional economics, Schultze had indicated useful directions, which I endorse. As a critic of the tendency of conventional economists to ignore the social system itself as a problem for analysis, Gurley has made clear the parochialism of much economic thought, immersed in technicism, blind to history.

But how far can conventional economic analysis go, knowing so little about the behavior of the public sector (not to mention its unhappy record of prediction with regard to the

private sector)? And what path shall we pursue to the construction of a new society, national or international? The answer to both questions, I submit, will require a profound change in the prevailing conception of economic theory itself—a change away from a view of theory as essentially predictive, basing its conclusions on the effects of supposedly known “laws” of behavior and technology, to a view of theory as essentially plan-oriented, making its primary task the exploration of possible paths toward a given objective. Such an “instrumental” (to use the term suggested by Adolph Lowe in *On Economic Knowledge*) conception of economics must no doubt rely to the fullest extent on whatever regularities, propensities, well-behaved functions, etc., it can muster, but ultimately it departs from the conventional conception in its willingness to search for, rather than to assume as given, the behavioral as well as institutional conditions required to reach social goals.

Indeed, under various guises, this instrumental reorientation of economics has already begun. In the reconstruction of society that Gurley envisages, economics must clearly take second place to social goal setting, clarifying the nature of the constraints, the sequence of processes, and the necessities for choice that all the moral fervor in the world will not cause to disappear, but not prescribing any particular course on *a priori* grounds. Meanwhile, in Schultze’s more limited delineation of appropriate social change, we find an explicit plea for “designing institutional structures and incentive systems which [will] automatically induce individual decision makers to move towards socially desirable ends.” I am not sure about the word “automatic,” but the spirit of using economics as a goal-achieving social science is certainly there.

The great question, of course, is what the goal is to be. One suspects that Messrs. Gurley and Schultze would not see eye to eye on this matter. But here, fortunately, the professional competence of the economist comes to an end, and he can do no more than take his place within the polity, to urge whatever goals—equality, freedom, growth, ecological balance or whatever—he seeks for society.

HENRY W. RIECKEN: As the only noneconomist at this table I feel not only outnumbered, but severely underequipped in technical knowledge to cope with the criticisms and counter-criticisms that my fellow discussants have presented so effectively. In view of the warm nature of some of the debate, I also have another feeling which I can best describe in terms more familiar to psychologists like myself. I feel a bit like a clinician in the private practice of family therapy on the occasion of his first interview with a new set of patients when he is attempting both to learn their private language and to understand the history and location of some of the conflicts they bring to him. Be that as it may, I want also to remark that I hope this marks the end of what could become an infinite regress. The economics panel reviewed the field. Professors Schultze and Gurley reviewed the panel report. Professors Heilbroner and Solow reviewed their reviews and I am joining in that effort. I doubt that I shall get the final word in this debate, but I promise you that I shall say no more about it after today.

The only thing that Gurley and Schultze have in common is a disappointment in the performance of their colleagues as their colleagues review economics. Their disappointments are quite different: Professor Gurley is angry whereas Charles Schultze is only sad. Both of them believe that the economics panel report did not do the job it should have done. In this respect, both critics put me in mind of a garage proprietor to whom I once brought an ancient Buick with the complaint that it wasn’t running at all well and a rather open-ended question as to what he thought should be done. His reply was not unlike that of Schultze and Gurley; he said, “Jack up the radiator cap and run a new car under it.”

Schultze is more tolerant than Gurley. He is willing to treat the report at one level as a “highly competent piece of work” seen as a “summary of the current status of economics as a technical discipline” and as a field “concerned with how society maximizes economic output subject to resource constraints in current social structures.”

Professor Gurley is less tolerant. He believes that the economics panel report speaks

for, and by, the establishment and I presume that my own account here of how the whole thing was generated and carried out only reinforces this conviction on his part. I find it hard to disagree with the details of Professor Gurley's argument, once I grant his initial premise, namely that the panel report should have been a revolutionary document, a document directed toward action, a critique rather than an exposition of a technically sophisticated social science which is devoted to the purpose of "maximizing economic output subject to resource constraints in current social structures," as Professor Schultze put the matter.

I find it hard to disagree with Professor Gurley's further exposition. I only disagree with his initial assumption. I do not believe that the BASS Report was intended to be, or should have attempted to be, the sort of radical critique that he calls for.

On the other hand, I think that the panel members, who addressed themselves to a somewhat different task than either Professors Gurley or Schultze call for, did correctly believe that they had been asked to address themselves to the somewhat different task. They had been asked to describe the state of the art, to give an account both of what economics had achieved, how it had contributed to society so far, and what some of the growing edges and potentialities of the field as an intellectual discipline might be. In this connection I am in full agreement with Professor Solow when he remarks that the panel was asked to review economics, not the economic system. After all, this exercise was put in the context of developing scientific disciplines, and the panel believed it was not being asked to address itself exclusively, or even primarily, to the use of economics as a problem-solving

tool. In fact, the charge to the panel was to produce a document which would advise the Congress, officials in the Executive branch, and individuals in the private sector as to what steps they should take and what they should avoid in developing a rational strategy for the strengthening and growth of economics as a social science.

The criticisms which we have heard this afternoon, then, may be more a reflection of the changes that have taken place in the temper of our times, the degree to which academic social scientists have become concerned with the application of their knowledge and the usefulness of their techniques than it is of anything else. Since this venture was conceived in 1966, there has been a substantial change in not only the political but also the intellectual atmosphere on American campuses and elsewhere in American life. What seemed like a sensible, indeed a very important task when it was conceived, now is being criticized as inadequate and perhaps wrongly directed. Perhaps it is inadequate. I certainly would not argue that the report of the panel or the overall report of the committee deals with all of the important issues regarding the social sciences today. Far from it. We did take a somewhat narrow, self-interested view and I think that was justified and is still sound. It is important to know how to improve the state of the social sciences and how to cultivate the sort of human talent and institutional resources that will improve the social sciences as intellectual enterprises.

It also is important to know how these intellectual enterprises may be turned to the greater service of mankind and it is that goal which our critics today told us was not achieved by the report.

TELEVISION: OLD THEORIES, CURRENT FACTS, AND FUTURE POLICIES

The Growth of Cable TV and Its Probable Impact on Over-The-Air Broadcasting*

By ROLLA EDWARD PARK

The RAND Corporation, Santa Monica, California

In January 1970, roughly 7.5 percent of U.S. television homes were cable subscribers, and this figure is increasing rapidly. Over the past decade, the number of subscribers has increased at an average annual rate of about 21 percent, while the number of cable systems has grown by 14 percent per year.

This remarkable growth raises a real dilemma for public policy. On the one hand, it holds out the promise of more diverse programming made possible, even promoted, by the ability of cable to carry a large number of signals. On the other, it poses a threat to over-the-air television broadcasting. When cable carries distant signals, it fragments the local audience, tending to reduce local station revenue. Reduced revenues may force broadcasters to reduce the quality of their programming, particularly local and public service programming. Some stations might even be forced off the air, reducing or eliminating the service available to cable non-subscribers. The loss to those viewers who are in areas not served by cable and to those who cannot afford the cable subscription

fee might be considerable. In other words, cable may provide more-diverse programming for its subscribers, but at the cost of reducing the amount of programming available to non-subscribers.

FCC Policy

In the face of this dilemma, Federal Communications Commission (FCC) policy has been to restrict cable growth in larger (top 100) markets, while allowing unrestricted growth in smaller markets, apparently opting for diversity for subscribers in the small markets and diversity for non-subscribers in the large markets. It is easy to formulate an argument of sorts in favor of such a choice. The largest 100 markets typically have stations affiliated with each of the three networks, and many have independent stations in addition. One can argue that this is "adequate," and that additional signals for cable subscribers are a superficial gain that does not justify jeopardizing continued adequate service to non-subscribers. On the other hand, most smaller markets lack complete network service. If this is judged to be inadequate, it is difficult to deny viewers the right to purchase adequate service via cable.

Some such argument may be a partial explanation of FCC policy toward cable. Additional explanation is provided by an

* This paper is based on a more detailed report, "Potential Impact of Cable Growth on Television Broadcasting," RAND R-587, October 1970, which was prepared under a grant from the Ford Foundation. All views and conclusions are my own, and should not be interpreted as the official opinion or policy of RAND or of the Ford Foundation.

examination of the historical roots of the policy.

Since the early days of radio, the FCC has sought to promote diversity of programming. In a free market, diversity requires (and to some extent follows from) competition among a large number of broadcast stations. In the case of radio, this is an approachable goal; for television, it is much more difficult. In the very high frequency (VHF) band originally allocated for television, most markets can have only two or three stations, and even the largest no more than seven, because of limited bandwidth and interference problems. To overcome this technical limit on the number of stations, the FCC has actively encouraged television broadcast in the ultra high frequency (UHF) band. In 1952, the Commission allocated a large chunk of the UHF spectrum for television. In 1964, to increase the potential UHF audience, it obtained passage of a law requiring all newly manufactured sets to have UHF as well as VHF tuners.

Thus, the FCC's goal of diversity has led to an organizational commitment to UHF broadcast. Nevertheless, UHF stations have, for the most part, continued to be marginal ventures. UHF stations without network affiliation, in particular, have had a rough time of it; only two out of 37 such stations reported a profit in 1968.

Most UHF stations are located in the larger markets; it is to these markets that the FCC looks for vindication of its promotion of UHF. This may help to explain FCC policy toward cable. By restricting cable growth in the top 100 markets, the FCC is shielding struggling UHF stations from the (supposed) harmful impact of cable.

Probable Impact of Cable

The Model

The FCC's policy toward cable has de-

veloped over the last several years in the absence of any detailed and comprehensive investigation of what the impact of unrestricted cable growth actually would be. This paper reports on an attempt to answer that question. My "impact model" is constructed from four separately estimated pieces:

(1) Certainly, the impact of cable depends on what portion of households can be expected to subscribe to cable service. Estimates of cable penetration ultimately to be expected are made by fitting a set of logistic growth curves to data on a fairly large sample of cable systems. Not surprisingly, ultimate cable penetration tends to be higher (a) for systems carrying a greater number of distant signals and (b) for systems operating in areas with fewer signals available locally over the air. For the classifications used, estimated ultimate penetration ranges from 29 to 60 percent. Rough calculations suggest an ultimate nationwide average penetration on the order of 40 to 45 percent of households.

(2) In estimating impact, it is of central importance to know what share of the audience will watch distant signals, and what share will continue to watch local ones. A method for estimating audience shares is developed, based on the following hypothesis: "Attractiveness" indices can be assigned to television signals so that for any set of signals, audience shares tend to be proportional to the indices. In the process of actually assigning such indices, several rough tests provide some support for the hypothesis. Of particular interest is an estimate, using 1968 data, that the attractiveness index for a network signal broadcast over UHF is only about one-half the index for the same signal broadcast by VHF, probably because of transmitter, antenna and tuner differences. Since the UHF handicap is wiped

out when the UHF station is carried on cable, this estimate provides one reason to expect that cable may help UHF stations, at least relative to local VHF stations.

(3) Taken together, (1) and (2) permit the calculation of station audience if distant signal carriage is specified. Relationships that aid in translating station audience into station revenue are estimated next. Of particular interest are two results indicating that an audience taken from local stations by distant signals is more valuable to the losing station than to the gaining station. First, the revenue-audience relationship is found to be curved in such a way that an additional household is worth less to a large station than to a small one. Second, distant audience is found to be worth less than closer audience. Since one result of cable growth is likely to be a decrease in the local audience of small stations and a corresponding increase in the distant audience of larger stations, both results indicate that total audience value will be decreased by cable growth.

(4) Next, the relation between local and public service programming, on the one hand, and station revenue, on the other, is investigated. In general, both the quantity and the quality of such programming are higher for stations with higher revenue. On the average, between 15 and 21 cents of each marginal revenue dollar is spent on local programming. This suggests that any adverse impact of cable on station revenue may well be reflected in decreased local programming.

These four pieces are put together to form the comprehensive impact model. A strong set of distant signals is assumed—signals from four very strong independents in the top 100 markets and three in the next 100, plus network signals sufficient to provide three-network service. Cable penetration is assumed to reach ultimate levels. The model provides detailed

estimates of station audience, revenue and local programming expenditure, with and without cable, in three different environments:

- 1960's environment, with UHF set penetration at November 1968 levels, and a UHF handicap due to antenna, transmitter, and tuner differences of about one-half, as estimated using 1968 data.
- 1970's environment, in which UHF set penetration is assumed to reach 100 percent, but the UHF handicap due to the other factors remains unchanged.
- 1980's environment, in which technological improvements are assumed to have eliminated the UHF handicap entirely.

The Results

Table 1, showing the impact of cable on revenue in the 1970's environment, is a sample of the results from the model. This table reflects only effects on the local audience. Stations carried by cable into distant markets have, in addition, a distant audience that also contributes to revenue. The contribution of the distant audience is discussed separately below.

Overall, station revenue (attributable to local audience) is reduced 18 percent by cable at its ultimate penetration, carrying the strong set of distant signals assumed. There is, however, considerable variation among markets and among different kinds of stations. Generally, stations in smaller markets are harder hit than those in larger markets. Those in large (top 50) markets lose, on the average, 15 percent of their without-cable revenue; those in small (fourth 50) markets lose 56 percent, on the average.

The reasons for stations in the smaller markets being harder hit are easy to see. Distant signals capture a larger share of

TABLE 1—PERCENTAGE CHANGE IN REVENUE DUE TO CABLE IN A 1970's ENVIRONMENT^a

Type of Station	Market Rank				
	1-50	51-100	101-150	151-200	1-200
Network VHF	-17	-24	-31	-55	-20
Network UHF	+12	-13	-15	(b)	-14
Independent VHF	-11	(b)	(b)	(b)	-11
Independent UHF	+20	+20	(b)	(b)	+19
All	-15	-23	-30	-56	-18

^a All figures reflect the effect on local audience only. Distant audience increases values in some cases.

^b Classifications with fewer than five stations are not reported in detail, but are included in the totals.

the local audience when competing with a smaller number of local signals. Also, cable penetration is expected to be higher in markets with fewer local signals.

There is also striking variation in how different kinds of stations are affected. Generally, UHF stations are less harmed—many are even benefited—by cable than are VHF stations. Network affiliated UHFs in the model lose, on the average, 14 percent of non-cable revenue, and the revenue of UHF independents actually rises 19 percent above its non-cable level.

The general reasons for this differential impact between UHF and VHF are clear. Over the cable, UHF stations are on an equal footing with VHF stations. It does not matter whether the cable subscriber lacks a UHF antenna or lives where UHF reception is poor. He gets UHF stations with the same click-stop tuning as VHF stations. The gain from achieving technical parity with VHF over the cable tends to offset, and in some cases more than offsets, the loss from audience fragmentation.

Among the UHF stations, the network affiliates are harmed by cable while the independents are helped. There are two explanations for this. First, the principal competition of independent UHFs is VHF stations, but many UHF network stations compete with one or two other network

UHFs. Independents thus have more to gain than network stations from achieving technical parity with VHF stations on the cable. Second, network UHFs are generally found in smaller markets than are independent UHFs, and the smaller markets are harder hit by cable growth.

The discussion above reflects only the effects of cable on local audience. In the model, the gain in distant audience exactly equals the loss in local audience. The distant audience has some value, tending to offset some of the revenue losses discussed. It seems likely that stations carried as distant signals will tend to be fairly large ones with strong programming, certainly larger on the average than the local stations whose audience they capture. Results on audience-revenue relationships in (3) above indicate that additional audience is worth less to large stations than to small ones, and distant audience is worth less than local audience. The combined effect is to make audience worth about half as much to the gaining station as to the losing station. The net overall revenue loss attributable to cable is then 9 percent; half of the 18 percent loss in revenue based on local audience is gained back by stations carried as distant signals.

If, as seems likely, distant signals are taken mostly from larger markets, the differential impact of cable in large and

small markets, apparent in the table, is accentuated. Large-market stations, which lose least in terms of local audience, stand to gain most in terms of distant audience.

Additional results, relating to audience, profit and local programming expenditure as well as to revenue in the 1960's, 1970's, and 1980's environments, are also discussed in my RAND report. Generally, the patterns that emerge in Table 1 and the discussion above are repeated in the additional results.

Conclusion

Results from the model indicate that concern over the potential impact of cable growth on television broadcasting is misdirected on several counts.

(1) Reduction in aggregate station revenue due to cable is perhaps not large enough to justify any great concern. Over-

all revenue loss due to cable is estimated to be about 9 percent. This loss is small enough so that it would be wiped out by one year's typical revenue growth.

(2) Stations in larger markets, now sheltered by FCC policy, would on average be little hurt by unrestricted cable growth.

(3) Stations in smaller markets, for which FCC policy now provides no protection, would suffer severe revenue reduction due to cable at ultimate penetration. Many might be forced either to discontinue service or to continue only as a satellite of a larger station.

(4) In the near term, say through the 1970's, non-network UHF stations (the objects of particular FCC concern) stand to gain substantially from cable growth, because cable puts them on the same technical footing as competing VHF stations.

Concentration of Control and the Price of Television Time*

By JOHN L. PETERMAN
University of Chicago Law School

The FCC has recently been considering the issue of "concentration of control" of the broadcast media. This issue stems in part from the belief that the present degree of concentration adversely affects competition in the markets for time. This implies that the distribution of ownership affects the price of time, or stated differently, the price of TV audiences.

The FCC is mainly concerned with the substantial control of TV licenses by newspapers and radio stations in the same market areas, and with the extent of group ownership of stations particularly in the largest fifty markets where group ownership is felt to be excessive. As a result, it began to exercise its right to question license transfers and renewals which perpetuated or increased "undue" concentration of ownership, and in its 1965 policy statement indicated its intent to

give greater weight to diversity of ownership when determining license grants through comparative hearings. Most concretely, the FCC has adopted a rule forbidding any party from holding more than one full-time broadcast license in any market. This rule did not require divestiture and under pressure from the Department of Justice, the FCC is now considering requiring licensees to reduce their holdings in any market to an AM-FM combination, a TV station or a newspaper. But in all these deliberations the underlying approach has been to assert that concentration is too great without specifying what adverse effects it has, or how changes in ownership would affect them.

I first consider group ownership. The fixed number of TV stations in a given market area might collude in order to restrict the supply of audiences or to facilitate price discrimination. I have found no evidence of discrimination in the sale of time.¹ Therefore, it is my view that collusion would only occur to restrict the supply of audiences. The question then becomes what difference would group ownership make, since in any given market, only one TV station can be owned by a single licensee. One could argue that group ownership reduces the costs of collusion and thus increases its likelihood. For example, group ownership increases the probability that in any market two or more firms meet in some other market. If collusion is more likely where an arrangement covers more

*I thank Professor R. H. Coase and Richard O. Zerbe for helpful comments. The sources of information for this study are as follows: the 1967 market rankings, the number of TV stations by market in 1967, and the prices of time were obtained from *Spot Television Rates and Data*, Apr. 15, 1967, vol. 49, No. 4; the number, location and circulation of newspaper firms from *Ayer Directory of Newspapers and Periodicals*, 1968; audience sizes from American Research Bureau, *Day-Part Television Audience Summary*, Feb./Mar., 1967; the ownership of stations from *Broadcasting Yearbook Issue*, 1968 and Hearings on S. 1312 Before the Subcomm. on Antitrust and Monopoly of the Senate Comm. on the Judiciary, 90th Cong., 2d Sess., pt. 7, at 3303-60; the number of radio stations from *Broadcasting*, Feb. 10, 1969, at 45-59. All measures of the number of TV stations exclude, whereas audience measures for each TV station include, owned booster and satellite facilities.

Because of the length of the original manuscript, it was cut by the editor.

¹ See, Peterman, "The Clorox Case and the Television Rate Structures," 11 *J. Law & Econ.* 321 (1968).

than one market, or if firms colluding in one market are more likely to collude should they meet elsewhere, then rates would be higher. The plausibility of this argument disappears on examining the facts. On the one hand, the extent to which group owners meet in more than one market is quite limited. On the other, prices do not appear to be affected in markets where cross ownerships occur. In fact there were only nine markets, or about six percent of all markets with more than one station, in which stations were operated by firms which met in at least one other market. These nine markets were all ranked as within the 100 largest TV markets. The average cost per unit of audience reached in the "cross markets" did not differ significantly² from the average cost in all other of the 100 largest markets or, for that matter, from the costs in all TV markets.

Perhaps the FCC believes that for important groups of advertisers different markets are close substitutes and that therefore collusion is unlikely to be profitable, or perhaps, because of the limited number and effect of cross meetings, that the discussion of group ownership frequently centers around a TV market which is defined by a larger region as the nation as a whole, or as a large subgroup of cities. A group firm may find an increase in rates on one or more of its stations unprofitable as a result of shifts in demand to other markets. If firms simultaneously agree on rates, then as a group they could possibly gain. But if collusion is not profitable in a more narrowly defined market, it is unlikely to be profitable in one which is more broadly defined. Since the maximum number of stations a firm can own is fixed by law, the more widely the market is defined the smaller relative to this market the firm becomes. In 1967, there were 92 group firms owning

54 percent of the TV stations in cities containing more than one station. Although in the largest 50 markets the number of group firms owning stations is only 55, and the proportion of stations they own 70 percent, the absolute number of firms would seem to preclude agreement. This view is reinforced by the relative size of group firms. There were 15 group firms (excluding the networks and their owned stations) sufficiently large to obtain above one percent of the national TV audience in March 1967. The largest firm held four percent of the audience; the average firm 1.9 percent. Including the network firms (each of which on average obtained seven percent of the total audience), the average for the 18 largest firms was only 2.7 percent of the total audience. Given these facts I would not expect group ownership to have an independent effect on the price of time.

We now turn to joint newspaper-TV ownership in the same market. If the demands for time and space confronting such a firm are competitive but not perfectly elastic, then in setting prices the interrelationship between demands may be taken into account. At the margin there is bound to be an effect which, if sufficiently large, would cause prices for time and space to differ from what they would be if the products were supplied independently.³ The Department of Justice and the FCC could argue that rates for time would be higher than without joint ownership. But the effect at the margin may be small and too costly to take into account. In effect the prices of time and space would be set independently. Which situation will be found cannot be determined by assertion. Given that a large proportion of the receipts of daily newspapers are derived from those who advertise little on TV, and

³ See, R. H. Coase, "Monopoly Pricing With Interrelated Costs and Demands," 13 *Economica*, N.S. 278 (Nov. 1946).

² Significance here and elsewhere is at the .05 level.

given the existence of other stations and competing media in most markets, a rise in the price of time may not cause a significant increase in demand for space so that the products are in fact priced independently.

Interrelationships in demand (and costs) may not, however, be the only reason why rates for time would differ in markets where joint ownership exists. Rates may be higher if joint ownership facilitates collusion between owners of competing media. But collusion has costs; it has probable gains, and the costs must be relatively lower or the gain relatively greater if joint ownership is to have an independent effect on the likelihood of collusion. The gains from collusion cannot be determined independently of the nature and extent of the interrelation in demand. If the demand for TV audiences is elastic but supply quite inelastic, the effect of collusion on the price of time could be extremely small.

I assume that no restrictions prevent buyers from choosing among stations or papers for desired quantities of advertising, i.e., purchases of time and space are not contingent upon the purchase of the other. I know of no examples of sales of time tied to space, or of the dual rate structures which such may require; nor am I aware of refusals to sell one product unless the buyer deals exclusively with the seller in the purchase of another. Such practices if they exist may cause the prices of time of the newspaper station to differ from other stations in the market. But in fact, I was unable to find any difference between the rates of stations which were newspaper owned compared to those which were not, once account was taken of the audiences reached by each TV station and of income in the stations' broadcast area. This is also true of the rates of stations which owned radio stations in the same market compared with

those which did not, and of the rates of stations which were or were not group owned. The latter two results would be expected, since in a given market area TV stations are excellent substitutes.

Much the same reasoning applies to radio ownership of TV stations in the same market. One cannot simply assert that effects exist. They may not because the interrelations in demand (and costs) are not worth taking into account, collusion may not be feasible or worthwhile, and the joint firm may have little influence on the supply of radio audiences, so that joint ownership has no independent effect on TV rates.

Evidence on the Effects of Ownership on Rates for Television Time

The above arguments imply that differences in ownership may affect rates for TV time. I now report the results of an investigation of whether in fact differences in rates exist. This is a necessary step though one which neither the FCC nor the Department of Justice felt required to undertake.

The prices used are the stations' 20-second announcement rates charged national advertisers. Most tests were also undertaken using alternatively the stations' network and national spot hour-rates. The results using hour rates are consistent with those using 20-second rates and therefore are not reported. Obviously advertisers purchase time only in order to reach audiences. The measures of audiences size (which are subsequently related to rates) are reported for each station in terms of the average reached (per quarter hour, from 6:30 to 10:00 or 7:30 to 11:00 p.m., depending on the time zone, over a two-week period). The audiences viewing any station vary in size: some programs are more popular than others, but in addition the proportion of total homes tuned to TV diminishes towards the fringes of the above daily

time periods. As a result, stations vary prices to account for the differences in the value of time. Therefore, the stations' 20-second rates (measured at the half-hour) have been averaged over the same period as that over which audiences were measured. Then for each station the average rate was reduced by the applicable discount for the purchase of three and five units per week. (The reason for discounting is not obvious and a word of explanation is required. As I have argued elsewhere, discounts were a means by which stations varied the prices of time to account for differences in its value.⁴ I would therefore expect discounted average prices to be more closely related to average audiences than would be the case if discounts were not taken into account.)

To examine the relation between audiences and rates for the purchase of one, three or five units per week, the average prices for each of 499 stations were viewed as a function of the average thousands of homes reached by the station and family income in each station's broadcast area. A regression was fitted from which the regression coefficients, *t* values, partial correlations and *R*²'s are listed below:

	Regression Coefficients	<i>t</i> values	Partial Correlations	<i>R</i> ²
20-second rate				
1 per week				
Homes	2.704	60.447	.938	.90
Income	.023	6.053	.263	
20-second rate				
3 per week				
Homes	2.640	74.465	.958	.93
Income	.019	6.313	.273	
20-second rate				
5 per week				
Homes	2.623	78.116	.961	.94
Income	.017	6.009	.260	

The discounted rates yield more highly significant relationships to homes actually

reached than nondiscounted ones. (This is consistent with the use of discounts to vary prices with changes in audience sizes and thus in the value of time.)

There are other possible measures of audiences. An obvious substitute would be the adults viewing each station. But the simple correlation coefficient between homes and adults reached (for 499 TV stations) is .998. Thus it matters little which is used. Only homes are reported in the remainder of this investigation. Aside from income, which in the above regression is a significant variable, it is not known whether other qualitative differences in audiences affect advertiser decisions and through them station rates. One such difference among stations would be in the average proportion of men to women viewers. But when this variable, along with homes and income, was regressed on average prices, its coefficient was not significant. There is no reason to assume that qualitative differences among audiences (which are not studied in further detail) would be systematically related to the ownership of TV stations.

Do prices differ in markets where newspaper firms jointly own TV stations? Each TV station's average price for five announcements per week was assumed to be a function of the homes reached, family income, and a dummy variable equal to one if the station is, or competes with, a newspaper-owned station, and to zero otherwise. A regression was fitted to this data for 499 TV stations. *R*² equaled .94. The regression coefficients, *t* values, and partial correlation coefficients are as follows:

	Regression Coefficients	<i>t</i> values	Partial Correlations
Homes (thousands)	2.624	77.385	.961
Income	.017	6.033	.262
Is or competes with a newspaper-owned station	-5.928	.922	-.041

⁴ Peterman, *supra* note 2, at 389-95.

That the coefficient of the dummy variable is not significant is consistent with the view that market prices per unit of audience do not increase as a result of newspaper ownership.

The newspaper-owned stations were next studied separately. Average discounted prices were regressed against homes reached, family income, the station's average share of TV audience, and the firm's share of the newspaper circulation either in (a) the city in which the station was located or alternatively in (b) the market area served by the TV station. In each instance the share of the newspaper is its circulation taken as a percent of the total circulation of daily newspapers published in the city or group of cities. The sum of the firm's share of the city newspaper circulation and its share of the TV audience was included as an independent variable along with homes and income in a separate regression. In total, each regression contained 90 observations. For each, R^2 equaled .90. The results are listed below:⁵

	Regression Coefficients	t values	Partial Correlations
Homes (thousands)	2.846	22.150	.923
Income	.016	1.773	.188
Percent Share of TV Station	.133	.327	.035
Percent Share of Paper (city)	-.520	1.082	-.117
Percent Share of Paper (market)	-.309	.915	-.098
Sum of Shares	-.242	.098	-.119

It might be expected that the greater the relative position of the firm in the TV market, in the newspaper market,

⁵ The coefficients for homes and income are given only for the regression containing the audience share of the TV station and the share of the (city) newspaper circulation.

or in both, the more likely there would be some effect on TV prices. This did not appear to be the case. The coefficients of each of the variables indicating the position of the joint firm in either or both markets were not significant. These results suggest that the demand interrelationships are not sufficiently important to take into account even if the position of the joint firm in the supply of space is large. Both products appear to be priced independently.

One further examination of newspaper ownership was made. The results are consistent with the above. The dependent variable is the homes (in thousands) reached by each TV station per dollar expenditure per 20-second announcement (valued at discounted rates for five units per week). The dependent variables are the number of competing TV stations, the number of radio stations in the TV market, the number of independent daily newspaper firms in the station's city, income, and a dummy variable (valued as previously) indicating whether or not the station is owned by a newspaper or competes with a newspaper-owned station. For 375 observations (all stations for which data was available), R^2 equaled .058. The results are listed below:

	Regression Coefficient	t value	Partial Correlation
Income	-.00002	2.032	-.105
Number of TV Stations	-.011	1.106	-.057
Number of Radio Stations	-.004	1.550	-.131
Number of Papers in city	.046	2.709	.139
Is or competes with a newspaper-owned station.	.006	.339	.018

The variable indicating whether or not the station is or competes with a newspaper-owned station is not significant. There is

virtually no relation between this variable and homes reached per dollar expenditure. Income is a significant variable (the negative sign indicating that increases in income are associated with decreases in homes reached per dollar expenditure). The number of independent daily newspaper firms in the city is also significant. The positive coefficient is consistent with the view that an increase in the number of newspaper firms tends to reduce rates for space, which in turn reduces the demand and thus rates for time. The number of TV stations per market does not appear to affect the cost of reaching audiences of given size. The main effect of increasing the number of stations per market seems to be to reduce the homes reached per station but not the average cost for given audience sizes.

I now turn to radio ownership by TV stations broadcasting in the same market area. In most markets containing two or more TV stations, at least one TV jointly owns a radio station. But between markets there is considerable variation in the proportion of the total number of radio stations jointly owned by TV firms. I assume for each market that this percent represents the degree of control over radio by TV stations. This is an imperfect measure, but data is not available for the receipts of individual radio stations, nor had I access to the audience shares. I tested whether differences in control over radio, so measured, affect prices of TV time. For each of 204 markets, the average discounted rates of each TV station are summed. These rates are assumed to be a function of the total homes reached in each market, of family income, and the percentage variable described above. Regression results are listed below. The R^2 of .97 is very high, reflecting elimination by summing of the small differences between stations in their prices relative to the audiences each actually obtained.

	Regression Coefficients	t value	Partial Correlations
Homes (thousands)	2.879	61.235	.974
Income	.037	2.930	.230
Percent of Radio owned by TV stations	-.907	1.302	-.126

The percentage variable is not significant, suggesting that increases in the control of the radio market does not affect rates for TV time. The regression was run separately for markets with exactly three TV stations. The results are essentially the same. In addition, for 51 markets containing only one TV station, discounted prices were regressed on homes reached, income, and a dummy variable equal to one if the TV jointly operates radio facilities in the same market and to zero otherwise. These are the smallest TV markets, and the relative position of the joint firm in both products is likely to be substantial; in addition, in more than 40 percent of the markets, the radio and TV stations were independently owned. The results are listed below. R^2 falls to .68, suggesting that in very small markets there are other factors affecting prices of time. But perhaps audience data is not reliable in very small markets.

	Regression Coefficients	t value	Partial Correlations
Homes (thousands)	1.619	9.878	.821
Income	.272	1.856	.194
TV owns Radio	-.707	.181	-.026

However, the coefficient of the ownership variable is not significant. This is consistent with ownership of radio by TV stations not affecting the prices of TV time.

I turn now to group ownership. A sample of 97 TV markets was selected. The sample contains markets with three stations plus all other markets in which three

stations obtained more than 90 percent of the audience. The markets were then divided into two groups—those which were and those which were not ranked within the 50 largest TV markets. These two groups differed considerably in the “degree” of group ownership of TV stations. In the larger markets the mean percent of the TV stations owned by group firms per market equaled 70, in the smaller markets 48. In the larger markets, the mean number of stations owned in total by the group firms located in each market equaled 8.84, in the smaller markets 4.77. The latter result is due largely to the fact that (relative to the number of markets) more group firms own stations in the larger markets. Thus in the larger markets, for a given number of stations, there are fewer independent firms. Then for each of the 97 markets the average homes reached per dollar expenditure (per 20-second announcement) was derived. The means of these values for the large market group equaled .32 thousand (assuming nondiscounted rates) and .34 thousand homes per dollar (at rates discounted for five units per week). For the 54 smaller markets, the means equaled .33 and .37 thousand homes per dollar. For all TV markets, the respective means equaled .31 and .36 thousand homes per dollar spent. The differences between the sample means and the means for all TV markets are not significantly different from zero. Thus there appears to be no difference in prices per unit of audience in the large compared to the small market group. But, in each of the two groups, there was considerable variation between markets in the degree of group ownership. Therefore, the sum of the discounted prices in each market was assumed to be a function of the sum of homes reached, family income, and alternatively the two measures respecting

group ownership. Thus two regressions were fitted to the data for each of the two market groups. For the larger markets (43 observations) R^2 equaled .95; for the smaller markets (54 observations) R^2 equaled .84. Listed below are the results for the two measures of group ownership.

	Regression Coefficient	t value	Partial Correlations
<i>54 Small Markets</i>			
% Group owned	.063	.178	.025
Total Stations group owned	1.800	.542	.076
<i>43 Large Markets</i>			
% Group owned	.107	.900	.143
Total Stations group owned	4.804	.552	.088

The coefficients of each of the variables are not significant. Differences in the degree of group ownership between markets does not appear to affect the prices of television time.

Conclusion

The evidence presented in this paper indicates that the present distribution of station ownership does not affect the prices of time to national advertisers. It is a simple matter for the FCC and the Antitrust Division to assert that important effects exist. It may also be a simple matter to base policy on these assertions. But on examination, the assertions may prove to be unfounded. Yet, the FCC and the Antitrust Division do not seem to feel obliged to make such an examination. The FCC and the Antitrust Division also assert that the existing pattern of ownership adversely affects programming. Would a detailed examination reveal that multiple-ownership does, in fact, have a significant adverse effect?

Program Duplication, Diversity, and Effective Viewer Choices: Some Empirical Findings*

By HARVEY J. LEVIN

Hofstra University and Center for Policy Research

In broadcasting, one standard by which industry and regulatory performance are increasingly evaluated is program diversity. The government hopes thereby to avoid any charge of dictating audience preferences while at the same time safeguarding those minority rights which the industry is sensitive to being told it has ignored. Program duplication and diversity assume special importance in an age when informed experts disagree on the precise program mix best-suited for democratic social-political processes. They are important also where the continued absence of demand prices for particular program outputs further complicates the task of regulation, as do current spectrum scarcities and the technical barriers to entry that result.

The wider the range of available program types, the greater the number of possible choices and the more varied the tastes and audiences accommodated. However, new TV entrants will tend to duplicate their incumbent rivals' programming so long as the market share they can thereby command will exceed what they could otherwise preempt with a new program type. According to P.

Steiner and J. Rothenberg, duplication is more likely to occur, the fewer the stations in the market at the outset, the greater the disparity between the sizes of the different listener groups and the smaller the disparity between the audience shares of competing stations.

Dramatic advances in the cable and satellite technologies, plus devices to impose per program viewing charges, now make possible greater diversity in the near and middle term. But how much diversity will materialize in fact, and how quickly? In seeking an answer, it makes sense to start by analyzing today's conventional TV broadcast programming, still a much neglected data source in a most complex empirical problem. Of three approaches examined below, the first has been widely utilized, the second was innovated two years ago in an industry-commissioned study of diversity, and the third, following Steiner, is offered for the first time here today.

I. Prime Time Shares of Major Program Types: An Exploratory Study

The major TV programming studies have normally sought to estimate the relative amount of broadcast time devoted to different program categories on the network showsheet, or on selected groupings and types of stations. This approach has serious limitations for any analysis of program diversity, but the basic information

*The ongoing research reported here is supported by the Augustus B. Weller Chair in Economics at Hofstra University and a grant from the National Science Foundation. I am greatly indebted for frequent talks about the regressions with Howard Kitt who also commented on the final draft, and to Richard Walko for invaluable help with all computer runs.

it generates still makes it a likely place to start.

Derived from data compiled from the 79 regional issues of *TV Guide* for the week of Feb. 25, 1967, e.g., my initial findings appear broadly consistent with those of studies by others in the mid-1950's. In Table 2, with changing numbers of commercial stations there is rough stability (except in the largest markets) in both the aggregate prime time share of three overall program categories and in the mean values of the specific types allocated to each of them. Superficially this pattern suggests growing duplication up to a point, insofar as the program mixes in markets with more stations resemble those with fewer stations. On the other hand, Table 1 shows that the programming of ETV and commercial stations decisively supplements (not duplicates) each other, irrespective of market size. Therefore ETV's entry into any commercial market should alter the program mix there away from the entertainment-types towards the informational ones. In Table 2 this shift actually occurs when ETV appears in an erstwhile three-station market which suggests greater diversity, sooner, than with new commercial entry.

One critical dilemma in this approach is that to report the full array of time-shares information for a substantial breakdown of program types, separately by number and type of station and market size, would produce a bewildering mass of results impossible to draw any but the most general conclusion from. Yet the mean values of prime time shares of selected groups of program types (assigned to a few overall classes) might well mask inter-category shifts as the number and types of stations vary. A more explicit analysis of viewer choices and some simplifying statistical construct are clearly needed. Fortunately the extensive classification of TV programs, the substance of

TABLE 1—COMPARATIVE PRIME TIME PROGRAM ANALYSIS, 150 MARKETS, FEB. 1967
(in proportions of 1.000)

	No. of Types	All Markets		Top 25 Mkts		Second 25 Mkts		Third 25 Mkts		Top 75 Mkts		Bottom 75 Mkts	
		ETV	Comm'l	ETV	Comm'l	ETV	Comm'l	ETV	Comm'l	ETV	Comm'l	ETV	Comm'l
Entertainment	11	.020	.072	.020	.069	.020	.074	.021	.072	.020	.071	.013	.073
Type I Info	5	.100	.027	.097	.027	.101	.025	.104	.028	.100	.028	.130	.027
Type II Info	4	.053	.008	.059	.011	.049	.007	.048	.006	.053	.008	.049	.004
						<i>Mean Proportions Within Overall Categories*</i>							
						<i>Aggregate Proportions of Three Overall Program Categories**</i>							
Entertainment	1	.255	.826	.250	.807	.268	.842	.266	.830	.259	.824	.168	.830
Type I Info	1	.585	.140	.557	.146	.581	.129	.641	.142	.580	.140	.666	.140
Type II Info	1	.159	.032	.192	.045	.150	.027	.092	.026	.159	.034	.164	.028

TABLE 2—PRIME TIME TV PROGRAM TIME SHARES BY NUMBER AND TYPE OF STATION AND MARKETS SIZE, FEB. 1967
(in proportions of 1.000)

No. of Sta. in Mkt.	No. of Mkts.		Entertainment		Type I Info		Type II Info		Entertainment		Type I Info		Type II Info	
	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV	Comm'l Only	Comm'l & ETV
1	54	53	.072	.072	.028	.028	.009	.009	.823	.823	.144	.145	.031	.031
2	22	19	.072	.072	.029	.030	.005	.008	.826	.819	.146	.153	.026	.027
3	51	34	.073	.072	.026	.028	.005	.008	.838	.825	.133	.143	.027	.030
4	15	25	.072	.064	.028	.038	.006	.014	.825	.744	.143	.206	.039	.049
5 & more	8	19	.067	.063	.028	.038	.012	.017	.792	.732	.149	.206	.058	.061

Notes on Program Coding Used in Tables 1 and 2

- (1) *The Code: Entertainment*: Adventure/Mystery/Police; Comedy; Variety; Westerns; Feature Film; Light Drama; Heavy Drama; Light Music; Heavy Music; Sports Events; Specials. *Type I Information*: News/Weather/Sports; Travel; Light Talk/Variety Info; Heavy Talk/Special Talks; TV course. *Type II Information*: Discussion/Debate/Interviews; Documentary/News Specials; Religious; Political/Editorials.
- (2) *Its Derivation*: Especially helpful were G. Steiner, *The People Look at Television* (1963), chs. 5-6, and App. C, Tables 10-11 (Master Program Codes); the typology used by A. C. Nielsen & Co.; and the very detailed program classification devised by the National Association of Educational Broadcasters in the mid-1950's. Pertinent points were also raised by the program departments of the major TV networks.
- (3) *Its Application*. The study was limited to the prime time programs (6-11 pm) of all stations (except satellites and foreign language) in the 75 largest and 75 smallest markets reported in all issues of *TV Guide*, Feb. 25-March 3, 1967. To minimize errors of judgment, programs were normally allocated according to official breakdowns in the Network Show sheets, the published descriptions in *TV Guide*, and the listings and classifications of Broadcast Measurement Bureau, *Series, Serials, and Packages—A TV Film Source Book*, Vol. 6, Issue #2 (1965-66), and Issue #2S (Fall 1966). In borderline cases, telephone inquiries were made. When in doubt programs were classified against the hypotheses being tested or, where deficiencies could not be rectified, simply classified as "not available". For analytic convenience all logs were subdivided into 15-minute units. All market categories refer to markets in sample only, and may not coincide with the actual distribution of TV markets.

* = calculated as follows:
$$\sum_{i=1}^n \frac{\text{Program Time Type}_i}{\text{Total Prime Time}}$$

n

** Aggregate proportions may not sum to 1.000 due to rounding.

most past studies, permits construction of an Index of Program Diversity which meets both requirements.

II. *An Index of Program Diversity*

Shortly after completion of my exploratory research, H. Land Associates, under commission by the National Association of Broadcasters, prepared an Index of Diversity based on a similar 20-point coding of logs for 188 TV stations located in a sample of 59 markets covered by *TV Guide* for the week of April 27, 1968. (See H. Land Assoc., *Television and the Wired City*, 1968, ch. 2). The Diversity Index was defined as a weighted mean rank of program categories in the market, using the number of broadcast hours as weights, such that $DI =$

$$\frac{\sum rh}{\sum h}$$

wherein "r is the rank (1-20) in the particular market for any program classification, determined by ordering categories according to total broadcast time devoted to each (during a sample week); and h is the number of half-hours of programming time devoted to the category." As thus conceived, low values of the Index would indicate concentration towards the most popular program types, and high values greater equality of time for the less popular types of greater interest to minority taste groups.

Among other things, Land Associates concluded that a rise in commercial stations from one to six (for the evening schedule) raised the Diversity Index 20%, though at a declining rate as each station entered the market. The Index rose 3.5% with a second station, but only 3% with a sixth station. In contrast, adding a single ETV outlet to a five-station commercial market raised the Index by

25% while adding one ETV to a two-station market raised the Index a full 41%. So at the margin, one ETV entrant was rated 8-12 times more potent in raising diversity than one more commercial station.

Though a valuable step forward, this approach too has important drawbacks. Conceivably low diversity as thus measured might be consistent with substantial viewer choices on an *hour-by-hour* basis. Suppose that each station in a two-station market carries the identical number of 30-minute units of four program types. Because the program rankings (and the weights) are identical for both stations, the Land Index would indicate zero diversity. But suppose, that the stations actually *staggered* the 30-minute periods during which they broadcast any common program type, such that viewers in each period during the week had a choice of two different program types. One could then say that viewers had the *maximum* effective choice, that is, two different program types for each 30-minute period, or 70 "actual" out of 70 "potential" choices.

Regulatory constraints will not normally induce profit-maximizing stations to stagger their programs in this way, and the Land Index may not be as erroneous as this extreme example suggests. Nevertheless, it seems essential to cast any Diversity Index in a form far more responsive to the pattern of hourly viewer choices.

III. *Analysis of Effective Program Options*

As a start, one could define a Coefficient of Viewer Choice as =

$$\frac{\sum P_o}{\sum P_t}$$

where $P_o =$ the actual number of program options that stations in each market generate during some series of arbitrarily de-

finer time periods; and where P_t = the maximum number of options viewers *could* have if each period were filled by each station with a program type different from that simultaneously broadcast by all other stations in the market. At the extreme, this coefficient would vary between the reciprocal of the number of stations in the market (perfect duplication) and one (perfect diversity).

As a practical matter, such a composite dependent variable poses the problem of downward bias with rising numbers of stations, due in part to the fixed number of program types used throughout, and in part to the smaller portion of prime time in which some stations in smaller markets broadcast at all. It may also increase the likelihood of spurious statistical results. These problems are more readily handled by defining "viewer choices" as the absolute number of prime time program differences in each market throughout the week, and the subsequent analysis finally employed that variable.

The results summarized in Tables 3 and 4 were actually derived from a modified

sample of 88 markets and 279 stations, left after the removal of all one-station markets (relevant to my initial study of prime time shares but not to viewer choices), and after further adjustments to insure the data comparability, accuracy and completeness needed for regressions. For convenience I converted my original 20-point code to coincide more closely with the one used by Land Associates.

A. Average Number of Program Options in Markets with Varying Numbers of Commercial Stations

Table 3 presents the unweighted mean values of the actual number of commercial program options, by number of stations in the market. For comparison I have added the comparable values from Land Associates. In both cases, the number of program options rises with rising numbers of stations, but the rate of increase of the options declines. As new stations enter, that is, there is a diminishing marginal increment in the absolute level of viewer choices.

TABLE 3—AVERAGE NUMBER OF PROGRAM OPTIONS BY NUMBER OF STATIONS IN MARKET
(Commercial Only)

No. of Comm'l Sta. in Mkt.	No. of Markets		No. of Stations		Average Number of Program Options			Percent Increases in Program Options			Change in No. of Sta. in Mkt.	Percent Increase
	Levin	Land	Levin	Land	Levin (eve.)	Land		Levin	Land			
						(eve.)	(full sched)		(eve.)	(full)		
1	—	12	—	12	1.00*	1.00	1.00					
2	19	12	38	24	1.74	1.72	1.71	74*	72	71	1 to 2	100
3	50	12	150	36	2.40	2.44	2.31	38	42	35	2 to 3	50
4	12	10	48	40	2.89	3.14	2.66	21	29	15	3 to 4	33
5	2	6	10	30	3.40	3.61	3.10	18	15	17	4 to 5	25
6	3	6	18	36	4.10	4.13	3.38	20	14	9	5 to 6	20
7	1		7		4.71			15			6 to 7	17
8	1		8		5.12			9			7 to 8	14
9											8 to 9	13
10		1		10		5.05	4.38		(5)	(4)	9 to 10	11
Total	88	59	279	188								

() = estimated from Land's curvilinear projection of the average number of options when stations rise from 9 to 10.

* = hypothetical value

Levin Program Code: Dramatic Adventure, Cartoons, Children, Situation Comedy, Drama, Quiz or Game, Popular Music, Fine Arts, Sports Events, Serials, Variety, Feature Film, News, Instructional (light), Instructional (heavy), Interview, Discussion & Debate, Political, Documentary & News Specials, Religious.

Comment: Distinction between "light" and "heavy" instructional programs adapted from Gary Steiner, *The People Look at Television*. All other categories correspond also to H. Land Assoc., *Television and the Wired City* (1968), see pp. 56-58.

TABLE 4—MULTIPLE REGRESSIONS ON VIEWER CHOICES IN 88 TELEVISION MARKETS, FEB. 1967
(Regression Coefficients and their Standard Errors)

Equation	Constant Term	Dependent Variable, X_1 , Total Choices									Adj. R^2
		X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	
1	76.9424	.0081* (.0043)		.8642 (.0565)	.5529 (.0261)	.5984 (.0432)	a	a	a	a	.96651
2	90.1182	.0073* (.0044)		.8678 (.0566)			.5587 (.0279)	.5671 (.0234)	a	a	.96644
3	89.8423	.0078* (.0044)		.8674 (.0567)					.5613 (.0299)	.5655 (.0230)	.96634
4	89.6549	.0077* (.0043)	.5648 (.0226)	.8670 (.0563)							.96634
Dependent Variable, X_2 , Commercial Choices											
5	77.1893	.0075* (.0037)		a	.5423 (.0236)	.5947 (.0392)	a	a	a	a	.94483
6	92.3780	.0067* (.0038)		a			.5481 (.0251)	.5594 (.0210)	a	a	.94422
7	92.0752	.0073* (.0037)		a					.5502 (.0270)	.5575 (.0207)	.94412
8	91.0142	.0069* (.0037)	.5576 (.0202)	a							.94412

Note: Of the 26 regressions run, equations are reported only for the best fit with the dependent variables. All coefficients significant at 1% level in two-tailed test, except those marked with asterisk which are significant at 10% level.

a= introduction of this variable lowered value of adjusted R^2 and coefficient was insignificant at 10% level.

As noted, the fixed number of program types makes for some downward bias, the probability of choice being smaller the more numerous the stations and vice versa. Nevertheless, my 20-point code uses the best-known program types, and the coded logs unquestionably represent the kind of TV output actually available today. Therefore, some of the bias presumably has substantive relevance to the real world of television. On one other practical count any distortion seems at best negligible, namely, the sheer detail of a 20-point code and the relatively few stations in my sample (one-seventh) operating in markets with more than four stations. It should in any case be noted that many of the categories are so similar that there is probably some offsetting bias in favor of choice.

B. Regressions on Viewer Choices

The full range of variables analyzed in Table 4 below are these:

Dependent Variables

X_1 = number of viewer choices on all stations in market (including ETV).

X_2 = number of viewer choices on all com-

mercial stations in market (excluding ETV).

Independent Variables

X_3 = ARB net weekly circulation in market, divided by number of quarter-hour program units broadcast by all commercial stations therein.

X_4 = number of quarter-hour program units broadcast by commercial stations.

X_5 = number of quarter-hour program units broadcast by ETV stations.

X_6 = number of quarter-hour units broadcast by independent stations.

X_7 = number of quarter-hour units broadcast by network affiliates or stations.

X_8 = number of quarter-hour units broadcast by single-owned stations.

X_9 = number of quarter-hour units broadcast by group-owned stations.

X_{10} = number of quarter-hour units broadcast by newspaper-owned stations.

X_{11} = number of quarter-hour units broadcast by non-newspaper stations.

It should be noted that X_3 divides net weekly circulation (a better predictor here than TV homes) by X_4 , a proxy for the number of commercial stations more accurate for my purposes than the actual number itself. Comparable proxies are used for

the other independent variables too, mainly because stations of all types do not in all market classes always broadcast weekly the same number of prime time quarter hour units. The regression coefficients indicate marginal *increments* of choice provided by the addition of commercial or ETV stations.

The principal findings are these:

- (1) In equation 4, the number of ETV program units (X_5) is substantially more important than commercial units (X_4) in raising the level of total viewer choices (X_1), holding constant market size (X_3). However in equation 8, ETV has no discernible impact on the number of commercial choices (X_2).
- (2) Equations 1 and 5 corroborate this in greater detail. Again, the number of ETV units is substantially more important than commercial broadcast units in explaining total choices however the commercial units are classified, whether as independent (X_6) or network (X_7), single owned (X_8) or group (X_9), newspaper (X_{10}) or non-newspaper (X_{11}).¹ But once more, ETV has no discernible impact on commercial program choice.²
- (3) In equations 1–3, though ETV's impact is in each case substantially greater than that of any other variable, the successive pairs of commercial subclassifications are also significant. Furthermore, the regression coefficient for network affiliates reveals a statistically greater impact than for independent stations. Though small, the difference between these coefficients in equation 1 is significant at the 1% level. In equation 2, on the other hand, the difference between the coefficients for group and single owners is significant at the 5% level only, whereas the difference between newspaper and nonnewspaper stations in equation 3 is statistically insignificant.
- (4) Looking next at commercial choices (X_2), once again network affiliates have a statistically greater impact than independents in equation 5, and group owners have a comparable edge over single owners in equation 6. But nonnewspaper stations now have a significant edge over newspaper stations in equation 7 (at the 5% level), contrasting results which required further inquiry. Regressions were therefore rerun for my 67 markets which had network affiliates only. The calculated differences between group and single owner coefficients were significant there in *neither* equation 2 *nor* 6 (even at the 10% level). However, nonnewspaper stations now made greater contributions to diversity than newspaper stations in both equations 3 and 7, with the differences in coefficients significant at the 1% level.

The bearing of these findings on broad-

¹ Severe space limitations prevented me from reporting all results in several stepwise multiple regressions. Hence the discussants could not know that when the regressions were run without reference to market size, each coefficient in Table 4 was actually a bit larger, though with a wider margin for ETV over commercial TV, and a narrower margin for network affiliates over independents. By eliminating market size differences as a factor in diversity (several proxies were used and significant), the relative impact of network affiliation, groupness and newspaper ownership, as well as ETV, could be measured more accurately. Presumably the larger stations in the larger markets, with more advertiser income, are on that count better able to sustain unprofitable news and public affairs programming, and hence to contribute to diversity as conceived here. Independent stations are also clustered in these top markets. Market size was therefore explicitly imposed as a constraint to keep the other coefficients conservative.

² In rerunning the regressions for a more homogeneous sample—those 67 of my markets which had no independent stations—introduction of ETV (X_5)

did raise the value of R^2 minutely in both equations 6 and 7, but the calculated ETV coefficients were only .0870 and .0909, respectively, significant at the 10% level.

cast regulatory policy can now be stated tentatively as follows:

- (1) Serious questions are raised here about the efficacy of new commercial entry as the sole or major instrument for widening viewer choices. An activated UHF table, or even CATV systems with 20-40 channels, are no *necessary* guarantee of diversity. However, the present economic-regulatory structure includes a dominant role for advertiser-supported national networking, a negligible role for subscriber systems, an existing technology for program production, transmission and interconnection, a resultant set of cost parameters and a given number and size distribution of TV program suppliers. Changes in any or all of these parameters could alter the above-cited relationships and many others reported in Table 4.
- (2) The decisively higher relative impact of ETV over commercial stations indicates a relative underinvestment in the former, other things equal, from the viewpoint of program diversity as conceived here. However, no measure of actual viewing has been included. Neither is it clear that with *comparable* numbers of commercial and ETV stations, the addition of another ETV in the sample markets would have had as large a relative impact. Nevertheless, ETV does have such an impact in *today's* market distribution of stations, and that, after all, is what regulators must work with for the near to middle term.
- (3) It follows also that, other things equal, a policy to strengthen public television is a far more effective way to widen available program options than a policy to diversify station ownership alone. Nevertheless, there may be overriding social and political reasons for licensee diversification.

Nor do my measures of program diversity rule out the potential dangers of group or newspaper owners injecting common editorial positions on their stations, or otherwise slanting news coverage. Only systematic content analysis can determine how real such potential abuses may be. Yet there is little evidence here that program diversity would decline significantly, if at all, were a policy of divestiture applied to newspaper or group owners and, indeed, in the *newspaper* case some small *increase* of diversity would appear to result.³ On both counts, regulators who opt for licensee diversification to diffuse social and political power can take some small comfort from the coefficients in Table 4.

- (4) As for innovative change, the ultimate impact on program duplication and diversity of improved microwave systems, domestic satellites, low-powered UHF stations, new multi-channel cable systems, or even a nationwide cable grid, will depend in part on the new cost levels which these transmission and interconnection technologies generate.⁴ But the relative size of the ETV coefficients in Table 4 underscores also the need to give the most painstaking attention to the organizational structure of those public, quasi-public and private entities which eventually own and operate the new communications modes—their sources of funding, and the explicit constraints to which the management of the new channel capacity is subject.

³ Even where statistically significant, the differences between the coefficients within these two pairs are obviously very small.

⁴ A suggestive review of such issues appears in L. L. Johnson, *The Future of Cable Television: Some Problems of Federal Regulation* (Rand Corp., RM-6199-FF Jan. 1970), pp. 26-40.

TV Program Diversity—New Evidence and Old Theories*

By EDWARD GREENBERG and HAROLD J. BARNETT
Washington University

Although many have criticized the lack of program diversity, few have defined the concept in an operational way for research or policy. Most of the economics literature and the major public policy issue are concerned with how diversity varies as a function of the number of channels available in an area.¹ The question of diversity is thus a special case of the more general problem of product and quality competition as performance variables and their relation to market structure.

I. Relationships between Diversity and Number of Channels

Two extreme views. A home within reach of one channel will be able to choose among thirty-five programming hours weekly from 6 to 11 P.M. If the number of channels is increased to two, and if the two channels never duplicate programming simultaneously, the choice increases to seventy hours, and so on. An increase in the number of channels, so long as there is no simultaneous duplication, by itself increases diversity and viewers' choices in a linear fashion. Let us define "diversity elasticity" as the proportional change in a measure of diversity divided by the pro-

portional change in the number of channels. Taking the number of different programs in each time slot as a measure of diversity yields an elasticity of one.

This simple-minded approach ignores the fact that additional channels may broadcast programs of the same general type as appear on other channels. But characteristics other than "general type" may be important to viewers. For example, a second channel broadcasting exactly the same set of programs as the first but in different periods may be quite significant to viewers. Still other possibly significant dimensions of diversity without change of "general type" are: independent, multiple sources of supply and open entry for new sources; variations in commentators, writers, and artists; quality improvements; small and large differences in content; and extent of advertising interruptions.

At the other extreme, one might say that all commercial television constitutes a type of product. David Blank, for example, differentiates between "... further choice within the more restricted universe of commercial television..." which he characterizes as an increase in the *quantity* of television signals, and *diversity*, which is television time devoted to "... serious drama, serious music, ballet, art, etc."² That is, television diversity is not increased when additional programs of the types that ordinarily appear on commer-

* This research was supported under a grant from the National Science Foundation. Thanks to Arthur Denzau for computational assistance.

¹ Many of the policy questions are discussed in Harold J. Barnett and Edward Greenberg, "A Proposal for Wired City Television," *Washington University Law Quarterly*, Vol. 1968, No. 1. (Winter 1968), 1-25 and "Regulating CATV Systems: An Analysis of FCC Policy and an Alternative," *Law and Contemporary Problems*, Vol. 34, No. 3 (Summer 1969), 562-85.

² David M. Blank, "The Quest for Quantity and Diversity in Television Programming," *American Economic Review*, Vol. LVI, No. 2 (May 1966), 448-56.

cial television are made available, and program diversity with respect to number of commercial channels has a zero elasticity. The virtue of this hypothesis is that it focuses on an aspect of diversity which is central to much of the public criticism of U.S. commercial television: the relative neglect of cultural and educational programs for minority tastes. This view ignores the diversity from increases in the volume of commercial offerings and may be unduly pessimistic about the possibility that the offerings of "... serious drama, serious music, ballet, art, etc." might be improved by commercial stations, under appropriate market conditions.

Program types. A number of authors have worked with program types as a way of avoiding the extremes discussed above.⁵ The models may be simplified as follows:

1. There are program types.
2. Audience preferences are distributed unevenly over these types.
3. Commercial stations aim at maximizing numbers of viewers, because advertisers pay for numbers.
4. A station will offer the same program types as other stations as long as his expected share of audience is larger than the audience he would obtain by choosing a different type.

The result is that certain kinds of programs will be presented frequently, and others will not be shown at all.

This persuasive doctrine implies a diversity elasticity between the extremes. It finds that program-type diversity is not a

linear function of station numbers: programs within the same type do not increase diversity, while programs of different types do increase diversity.

The problem is to define the central variable, "program types." The definitions should be based on consumer preferences,⁴ permit econometric analysis of the effects on diversity as we increase numbers of stations, and permit us to appraise the effects on diversity of alternative organizations of the television industry. But such construction of program types is extraordinarily difficult, with the result that the theory has not been successfully applied in an empirical study.

If we can specify program types within which we can reasonably believe that variations of quality, advertising, timing, author, performers, suppliers, and degree of competition do not contribute to diversity, then we can define diversity as differences in program type and the "program-types" model can be applied. We will be able to judge how many stations are too few for the desired level of diversity. But, *until* we can so define program types, we must be content with the important insight that program diversity is likely to be less with "small" numbers of suppliers than with "large."⁵

II. Published Empirical Work

*Herman Land Associates-Frederic Stuart.*⁶ This study attempts to examine

⁴This point is made explicitly by Steiner, *Quarterly Journal of Economics*, Vol. LXVI, No. 2 and by Kurt Lang, "Areas of Radio Preferences: A Preliminary Inquiry," *Journal of Applied Psychology*, Vol. 41, No. 1 (1957), 7-14.

⁵Another approach to broadcaster behavior which obtains similar conclusions may be found in Stanley M. Besen and Ronald Soligo, "Viewer Preferences, Market Structure, and Economic Efficiency in the Television Broadcasting Industry" (Mimeo. November 1969).

⁶Herman W. Land Associates, Inc., *Television and the Wired City* (Washington: National Association of Broadcasters, 1968). The index discussed in this section is attributed to Dr. Fredric Stuart on p. 2.

⁵For more detail see: Peter O. Steiner, "Program Patterns and Preferences and the Workability of Competition in Radio Broadcasting," *Quarterly Journal of Economics*, Vol. LXVI, No. 2 (May 1952), 194-223; Jerome Rothenberg, "Consumer Sovereignty and Economics of TV Programming," *Studies in Public Communications*, No. 4 (Autumn 1962), 45-54; Peter Wiles, "Pilkington and the Theory of Value," *Economic Journal*, Vol. 73 (June 1963), 183-200; and John J. McGowan, "Competition, Regulation, and Performance in Television Broadcasting," *Washington University Law Quarterly*, Vol. 1967, No. 4 (Fall 1967), 499-520.

what happens to diversity when markets with different numbers of commercial and educational channels are compared. Their results imply diversity elasticities, from one to ten stations, as follows:

Diversity index elasticity, day and evening programs—.005

Diversity index elasticity, evening programs only—.04

The Land-Stuart results thus yield elasticities close to zero; indeed, the index frequently falls as station numbers are increased. The reason for this odd outcome lies in the way diversity is measured; the variable does not utilize program types in the sense described above. Rather, it measures the similarity of the mean of the *composition* of program types among commercial stations, based on an arbitrary scaling and weighting scheme.

The Land-Stuart data before the erroneous scaling and combination into an index are potentially more interesting relative to the "program types" models. On the average, in half-hour time periods, days and evenings, the number of Land-Stuart program types rises from 1 to 4.38 (diversity elasticity = .4) as commercial stations increase from one to ten, and from 1 to 6.05 (elasticity = .56) in evenings only. These data would give support to the program-type models if the program types were well-conceived relative to consumer preferences, which the program-types models require. This is not the case. Land-Stuart types seem to be industry categories based on historical accident, rough and ready casual description, or source of supply. It is not clear to us why presentation of Oliver's "Hamlet," a Marx brothers film, and "High Noon" should be defined as no diversity because all are "Feature Films," while the addition of "I Love Lucy" ("situation comedy"?) and "Bonanza" ("serial melodrama"?) constitute a double increment of diversity.

*T. M. Jones.*⁷ A similar study is concerned with diversity in films as a function of the number of theaters in a city. Jones applied Land-Stuart's method of constructing a diversity index to films shown at theaters in Illinois and Indiana towns during a five-month period, and obtained results similar to theirs.

For example, residents of Indianapolis had their choice of 290 different films over the five-month period, not including "art" theaters, while those in the towns with one theater had an average choice of 42 different films. Yet, the Land-Stuart type of diversity index moves only from 1.85 to 2.00. The highest diversity index is in cities with six theaters, which show only a third as many films as the largest city and fewer categories of films.

*Lang.*⁸ A study by Lang of types of radio programs makes a serious effort to ascertain program types as they appear in the minds of consumers, and incidentally shows that conventional industry categories do not adequately represent these. Data were obtained from interviews in 497 Canadian households; respondents were asked to list their five most favored radio programs. From these responses, an overlap statistic was computed for each pair of programs mentioned. The observed overlap is compared with the overlap to be expected if people favoring program *a* were as likely to favor program *b* as by chance in the overall sample.

Grouping programs according to whether there exists a higher ratio than would be expected by chance, he identifies a majority taste, and a number of minority clusters. The groups cut across conventional program categories.

⁷ Thomas M. Jones, "Diversity in the Movie Theater Industry Relative to National Wired Television" (Mimeo., undated). Mr. Jones displayed considerable ingenuity and effort in preparing his study at the University of Illinois while an undergraduate.

⁸ See Lang, *Journal of Applied Psychology*, Vol. 41, No. 1.

*Goodhart-Ehrenburg.*⁹ Employing methodology similar to that used by Lang, Goodhart and Ehrenburg claim that, to a good first approximation, the proportion of people viewing two programs depends on the ratings of each program, irrespective of "program types," on conventional industry definitions. From this they conclude that there are no program types. But their findings may as well be interpreted as demonstrating only that conventional industry categories are not appropriate types for analyzing viewer behavior.

Other work. Factor analysis of correlations between program viewing habits has been undertaken in several studies; however, the results are not readily interpreted and there appear to be deficiencies in the model specification.¹⁰ A method used in a study of Japanese television audiences is promising—canonical correlations were computed between matrices of 0-1 variables representing preference types and conventional industry program categories.¹¹ Replacing categories by individual programs might yield information on the structure of program types. The important feature of this research, and that conducted by Lang, and others mentioned above, is the attempt to utilize viewer behavior to construct program types which might be used to measure diversity.

III. New Evidence

As an illustration of the way in which viewer behavior might be utilized in research on diversity, we analyzed share of

⁹ G. J. Goodhart and A. S. C. Ehrenberg, "Duplication of Television Viewing Between and Within Channels," *Journal of Marketing Research*, Vol. VI (May 1969), 169-78.

¹⁰ See for example, Charles E. Swanson, "The Frequency Structure of Television and Magazines," *Journal of Advertising Research*, Vol. 7, No. 2 (June 1967), 8-14.

¹¹ Meiko Sugiyama (Hori), "Multi-Dimensional Analysis of Radio and TV Program Category Preferences," Public Opinion Research Institute, Japan Broadcasting Corporation (Undated).

audience data for network feature films broadcast during 1968.¹² From Table 1, we note that the share of audience obtained by the various films varies considerably, belying the Land-Stuart suggestion that films "... consist of minor variation in programming which would satisfy substantially the same audience preference..."

Two interesting questions may be explored with the data:

1. Does significant variance remain when such obvious factors as month, network, and day are taken into account?
2. Does film type, as defined by standard film industry categories, add to the explanation of viewing behavior?

The results of a multiple regression analysis which controls for these variables follow:

Month—A set of dummy variables representing individual months proved statistically significant, indicating that the share obtained by films varies systematically through the year.

Network-day—The effect of these two variables is somewhat confounded because each network regularly broadcasts films on particular days. The combined effect of the two variables is highly significant.

¹² These data are taken from a Special Report, "What Happened on 1,001 Movie Nights," *Broadcasting*, November 3, 1969, 61-80.

TABLE 1—SHARE OF AUDIENCE OBTAINED BY NETWORK FEATURE FILMS 1968

Share of viewing audience (percent)	Frequency
19 and under	3
20.0-24.9	29
25.0-29.9	49
30.0-34.9	78
35.0-39.9	64
40.0-44.9	30
45 and over	9
Total	262

Source: *Broadcasting*, November 3, 1969.

Color—A dummy variable for color films indicates a positive, but insignificant effect.

Age of film—Dummy variables for films less than five years old, between five and ten years old, and more than ten years old, proved to be insignificant.

First Television showing—This variable was highly significant; the audience share of a film not being shown for the first time was 3.2 percentage points lower.

Type—Film type, entered as a set of dummy variables representing seven types according to industry definitions was not significant. On an individual basis, only the coefficient for Westerns was significant.

In summary, the results of the regression indicate that:

1. While a significant proportion of the variance in shares can be explained by month and day of week, which is likely to represent habitual viewing, the substantial amount of unexplained variance suggests that "characteristics" of the individual film are important.

2. Of film characteristics readily available for examination (color, age, "type," and first showing) only a variable representing first showing made any difference statistically.

3. The R^2 of .391 when all variables are included indicates reasonably high explanatory power for the variables, but also considerable variation due to film characteristics not accounted for by the above variables. This suggests that it is improper to consider films as being perfect substitutes for each other.

IV. Conclusions

With respect to a major policy issue, whether the FCC should take actions which would greatly increase channel availability, we believe that:

1. The number of different television offerings available during a given time pe-

riod is a most important measure of diversity and choice.

2. Program characteristics other than "type" are significant for viewers.

3. To the extent it is useful to classify programs into "types," these should be based on audience behavior. Even when so defined, however, choice within types may be as valuable to many viewers as diversity across types.

4. More open access for suppliers of television offerings is valuable for economic, social, and political welfare.

5. As evidenced by the rapid expansion in CATV subscriptions, a recent study by McGowan and Peck,¹³ various pay-TV experiments,¹⁴ and Park's work,¹⁵ the public values additional choices even within the customary range of television fare. It is willing to pay fairly substantial amounts to receive additional programming of the same "general type" as those already being received without charge.

As has been stated in previous theory, certain program types or variants within types (serious drama and music, for example) may lack popular appeal and therefore be seldom shown on commercial television systems. In addition to policies designed to lower costs and increase availability of channels, society may properly delineate a class of programs to be considered "merit goods," to be supported by license fees and access to advertising revenue, as in European public broadcasting, or by subsidy from tax funds.¹⁶

¹³ John J. McGowan and Merton J. Peck, "Estimating Consumers' Valuation of Additional Television Programming from CATV Data" (Mimeo., undated).

¹⁴ Evidence on the pay-TV experiments is cited in Blank, *American Economic Review*, Vol. LVI, No. 2.

¹⁵ R. Edward Park, "Potential Impact of Cable Growth on Television Broadcasting," Rand R-587-FF, October 1970.

¹⁶ See Ronald H. Coase, *British Broadcasting: A Study in Monopoly* (Cambridge: Harvard University Press, 1950); Asa Briggs, *History of Broadcasting in the U.K.* (London: Oxford U. Press, 1961, 1965); and Harold J. Barnett, "Resistance to Wired City," *Essays in Honor of E. M. Hoover* (forthcoming).

DISCUSSION

JOHN J. MCGOWAN AND MERTON J. PECK: The papers presented today illustrate that economists can make substantial contributions to enlightened discussion of performance and regulation in television broadcasting. Common to these four seemingly diverse papers is an emphasis on quantitative analysis of important policy issues in television broadcasting. This emphasis itself is a significant step forward. All too often, policy discussions have been generalized criticism of television's performance and qualitative assertions about the effects of industry practices. Beyond this, the papers illustrate how the concepts of microeconomics can be combined with quantitative testing to examine the performance of this industry and the rationality of present policy.

Professor Peterman demonstrates quite conclusively that concentration and ownership patterns have little influence on local station advertising rates, if those rates are converted into the standard industry way of expressing prices—cost per thousand viewers. This very process, however, nets out the most important aspect of product differentiation—audience size—so that comparison of price per thousand is, in effect, price comparison of undifferentiated products.

With undifferentiated products, price uniformity is to be expected, regardless of market structure. Hence it would be consistent with Peterman's findings to regard the television advertising market as either oligopolistic or purely competitive. In each case, an undifferentiated product insures price uniformity—just as, for example, the price of wheat and steel are both uniform among their sellers. But this is not the same thing as saying that vastly different market structure and seller concentration do not affect the price level and pattern of price changes. To show that concentration does not affect the price of television time requires some investigation of the dynamics of price changes rather than an analysis of price differences at one moment of time.

How this would be done with the same degree of quantitative precision that Professor Peterman brings to price differences at one moment of time is not easy to visualize. We do

have the strong suspicion that the kind of ownership variables that the FCC has traditionally been concerned with and that Professor Peterman investigates—cross ownership between radio and television, newspaper ownership and group ownership—are not as important for price behavior as is the role of the networks. The three networks own five stations each—all in large markets—and have affiliations with more than 500 of the nation's 700 commercial stations. Each network faces its two rivals in almost all markets. It may well be that these three oligopolists largely determine advertising rates.

It should be noted that Peterman's results can be viewed as consistent with a pricing process that assigns a key role to the networks. Peterman's price variable is individual station's rates for national advertisers. To some degree, national spot advertising and network advertising are substitutes, though advertisers do in fact often view them as complements. The essential point is, however, that there is probably a high degree of cross elasticity of demand between national spot and network advertising. With the network oligopolists setting similar prices per thousand, the link between the national spot and network advertising markets may partially explain the similarity of price per thousand among local stations in different markets. Peterman's finding of a high correlation between national spot and network station rates also supports this view of price making.

As Peterman recognizes, the major concern with ownership patterns is not with their impact on the advertising market. We think the important concern may be their impact on the content of local news presentation and political processes in specific localities. Unlike most prime time programming, local news is produced by the local station management and it is a commonplace that the ownership of several stations or perhaps better yet some combination of the local newspaper, radio and television station tends to insure that telephone calls to Washington, the State house, and City hall are returned promptly. Exactly how media ownership influences the political

process is still obscure. FCC policy regarding ownership concentration, however, appears more based on noneconomic considerations than on its impact on competition in the advertising market.

Clearly the whole issue of ownership patterns needs more investigation as to both the current trends and how they affect both economic and noneconomic behavior. As long as the local station markets are oligopolistic, one suspects individual managements will have considerable degrees of freedom, and how they use that freedom could be influenced by station ownership. Yet adding more channels through CATV, more competitive supply conditions for programs, more ETV stations may have greater consequences for dispersion of the control of media than concern with the concentration of local station ownership.

Both the Levin and Greenberg-Barnett papers follow the lead of Peter Steiner's seminal article, as the authors acknowledge. Professor Steiner's article advanced the proposition that the extent of program diversity, i.e. number of program types broadcast per period, would depend upon the number of competing broadcasters and the distribution of viewer preferences over program types. Steiner argued that unless the potential audience for various program types was quite similar, expansion of the number of broadcasters would tend to result in duplication of existing program fare rather than the introduction of new program types. Since Steiner assumed away the possibility that viewers might have preferences within program types, total viewer satisfaction could be increased only if additional stations broadcast previously unavailable program types. Such an assumption is, of course, an oversimplification since viewers also have preferences among program types, a point to which we shall return.

Professor Levin estimates the extent to which increases in total broadcast hours are allocated between existing program types and new program types within any quarter hour broadcast unit. He also determines how that allocation is related to station ownership characteristics. For this purpose he employs a dependent variable called total viewer options. The dimension of this variable is implicitly

quarter hours, since it is obtained by: (1) determining the number of program types in each quarter hour, and (2) summing those numbers for all the prime-time quarter hours in a sample week. This measure can be considered as the effective quarter hours, since the underlying assumption is that simultaneously broadcast programs of the same type are perfect substitutes. Effective quarter hours differ from total hours to the extent duplication occurs.

Professor Levin finds that, on average, adding a quarter-hour broadcast by a commercial station increases effective quarter hours somewhat less than 60 percent. If no duplication occurred, the increase in effective quarter hours, of course, would equal the increase in broadcast quarter hours. Thus, the implication is that the average degree of duplication among commercial stations is more than 40 percent. On the other hand, additional quarter hours broadcast by an ETV station apparently increase effective hours by almost 87 percent, implying such stations duplicate commercial programming only 13 percent of the time. Levin, therefore concludes that expansion of ETV stations is a more effective way of providing diversity than is expansion of commercial stations.

We have problems both with Levin's statistical procedures and with his measure of diversity. Levin finds all of his coefficients on numbers of quarter hours to be significantly different from zero. Although it may seem a minor point, we suggest that the null hypothesis against which he should be testing is that the coefficients are in fact unity, as they would be if no duplication occurred. This would apparently have no effect on his conclusions since the reported coefficients and standard errors are such that the null hypothesis could be rejected at accepted levels of significance. However, the results might be otherwise if, as we believe is likely, Levin's procedure imparts a downward bias to some of his coefficients.

Our suspicions of bias are raised by Levin's variable, X_3 , defined as net weekly circulation (NWC) divided by commercial station quarter hours. While Levin indicates that this variable was chosen over alternative measures because it gave a better fit to the data, he pro-

vides no explanation of why market size itself should be a determinant of the extent of diversity. Nor is the role of market size evident from the original Steiner formulation of the problem. Steiner's analysis emphasized the *distribution* of viewer preferences, not the absolute magnitude of potential audiences. Whatever the rationale for including a market size variable, we suggest the measure used should be chosen on the grounds of its suitability as a proxy measure for the hypothesized effect and not on goodness of fit criteria.

There are other serious questions raised by X_3 , Levin's market size variable. As indicated above, the dimensions of X_3 are NWC, therefore, homes per commercial quarter-hour broadcast. NWC is highly, though not perfectly, correlated with the number of TV-homes in a market. Numbers of commercial quarter hours is also correlated with number of TV-homes though less closely than is NWC. We would, therefore, expect NWC per quarter hour of commercial broadcasting to be positively correlated with variables measuring quarter hours of commercial broadcasting. Our suspicion then is that the variable X_3 is absorbing some of the effect of additional commercial broadcasting on diversity. (Note that one cannot correct for this effect, if it is believed to be present, by simply adding the coefficient of X_3 to the coefficients on the commercial broadcasting variables because the dimension of the coefficient of X_3 is effective quarter hours per home while the dimensions of the other variables are effective quarter hours per broadcast quarter hour). The same possibility of downward bias does not arise with respect to the coefficient on ETV broadcasting because only commercial quarter hours were apparently used in constructing X_3 . Thus we suspect Levin underestimates the impact of commercial broadcasting relative to ETV broadcasting on effective viewing options.

These suspicions are heightened upon further examination of his estimates. As noted, they imply an average degree of duplication among commercial stations in excess of 40 percent regardless of whether the stations are network affiliates or independents. Though there is admittedly a trend toward increasing frequency of deletion or delay of network pro-

grams, affiliates still rely predominantly on network programs during prime time. One would expect then that the degree of duplication among affiliates would closely parallel the degree of duplication among network schedules. Yet, using program categories similar to Levin's the extent of duplication in prime time network programming is much closer to 20 than 40 percent. Indeed, Levin's estimates of average number of options per program unit reported in his Table 3 indicate 1.74 in a two-station market and 2.40 in a three-station market. Implied in these estimates is an average rate of duplication of $(2.00-1.74)/2.00 = .13$ in two-station markets and of $(3.00-2.40)/3.00 = .20$ in three-station markets.

These calculations suggest to us that the "true" value of the marginal impact on diversity of additional commercial, particularly affiliated station, quarter hours may be substantially larger than Levin estimates and may, in fact, be quite similar to the magnitude he finds for ETV broadcasting.

If this is the case, rather different and more tenuous arguments must be employed to find that ETV expansion is superior to commercial expansion in producing diversity. Most obviously, a different conception of diversity will be required, and this is exactly the question dealt with by Greenberg and Barnett.

Their basic point is that if diversity is fundamentally a measure of the extent to which viewer preferences are satisfied, the diversity index must be solidly founded on knowledge of those preferences. Greenberg and Barnett provide cogent arguments and convincing evidence that various diversity indices based on conventional program categories are seriously deficient in this respect. Their finding that type of movie is of little use in explaining the wide variation in movie audiences agrees with our own results. The program type variables also add virtually nothing to the explanation of variances in audience shares for syndicated programming (first-run and off-network shows) shown in prime time, after standardization for numbers of viewing options available and for VHF-UHF differences.

As Greenberg and Barnett point out, findings such as these indicate that industry categories and measures of diversity based on

them have little correspondence with relevant viewing options as consumers perceive them. In addition, these findings and others which they cite indicate that viewers place a high value on additional choice within traditionally defined program categories. Thus, it may be a serious error to base any policy decisions on an analysis which uses a measure of diversity that places a zero value on additional choice within standard program types. A Levin-type diversity index does just that.

Greenberg and Barnett urge policy makers not to ignore the potential benefits of expanded channel availability independent of any contribution to diversity and urge researchers to develop more refined concepts and indices of diversity. We strongly support their recommendation to policy makers. We are skeptical of their recommendation to researchers, for we suspect that such a venture will be at least as unrewarding as the efforts of the past twenty years to devise improved measures of industrial concentration. Furthermore, we think that a different approach to diversity can provide more immediate insights into the response of industry performance to increased channel availability.

Both Levin and Greenberg-Barnett recognize that the most widely perceived deficiency of present day commercial broadcasting is its failure to cater to minority tastes. The implication of Levin's results is that this failure derives from a distribution of viewer preferences which is heavily weighted toward existing program types and from uniform, zero-pricing of television programs in line with the original Steiner model. What both papers neglect to mention is that the profit maximizing calculations of broadcasters, which determine their program offerings, depend not only on viewer preferences, but also on program costs. Nor do they acknowledge that independent stations do in fact cater to minority audiences. For example, we have estimated that syndicated programs broadcast by a VHF station in prime time in a market containing three affiliates and one independent achieve, on average, an audience of five percent of total television homes. There is no significant difference between the average audience shares of first run syndications and off-network shows. In con-

trast, network programs attract, on average, over 18 percent of total television homes in prime time. Since independent stations are the predominant purveyors of syndicated programming in prime time, they are catering to minority audiences at least relative to networks and their affiliates. To be sure, the minorities to which these stations cater are comprised of individuals who desire more choice within the popular program types and are not the same minorities which Levin and Greenberg-Barnett apparently have in mind.

Their minorities want programs with high cultural or informational content. Whether these high-culture minorities should take precedence over the more prosaic minority tastes cannot be answered by resort to indexes of diversity. There is no way policy makers can escape a determination of the costs and benefits, both public and private, generated by catering to various minority tastes if they wish to weigh expansion of commercial broadcasting against expansion of public broadcasting.

Fortunately our responsibility today is not that of policy maker, so we can and will sidestep this difficult question. On the other hand, we think we can show how to answer a different question relevant to the policy maker's problem. We assume that commercial stations would provide programming which caters to high-culture minorities if it were profitable to do so. Expanding channel availability in the context of a Steiner model is supposed to bring this result about by decreasing the profitability of existing programming relative to the profitability of introducing new program types. Our question then is what amount of expansion of commercial broadcasting would be required in order to make it profitable for such stations to broadcast high-culture, minority-taste programming. Space does not permit a full treatment of this question. We wish only to indicate how existing data can be used to provide an answer.

High-culture programming might be provided either via networks or directly by local stations on a syndicated basis. In either case, cost data on network programming in the 1968 season indicate that production costs for high-culture programming which did not employ established stars extracting high monop-

oly rents, would amount to 85 to 90 thousand dollars per half-hour episode. One useful question to answer is what audience size is required in order for a network to break even on such a program at this cost level?

In 1968 the three networks had net revenues from time sales plus program and talent charges after deduction of sales commissions and compensation to affiliates of \$963.3 million. Others have estimated that roughly two-thirds of network revenue is generated by sales in prime time, so that we may place total prime-time net revenue at \$642 million. The average prime-time audience in winter months for the three networks together amounted to 55.5 percent of the estimated 57.5 million TV-homes in the country. Since summertime viewing tends to be about two-thirds of wintertime viewing, average prime time audiences per half hour amount to almost 46 percent of total TV-homes. Assuming 2,555 ($= 7 \times 365$) prime-time half hours per year, indicated network net revenue is \$9.50 per half hour per thousand viewing homes. At a cost of 85 thousand dollars per half-hour episode, cumulative audience over the program's network showings must amount to almost 9.0 million homes. Assuming two network showings, as is the case with most network programs, the program would need to obtain a rating of about 9.5 on its first winter month showing and a rating of 6 on its second showing. Since these ratings have been calculated on the assumption that the program need only obtain sufficient revenue to recoup program production costs, it would seem safe to assume that networks would never provide programs with expected ratings of less than 10 on their first prime-time showing.

Current ETV programming obtains average ratings of about one. To be sure, these low ratings can be partially explained by the fact that most ETV stations are UHF stations. It is also alleged that the standard rating procedures underestimate ETV audiences. Finally, it is usually argued that ETV ratings are low, in part, because lack of funds prevents high-quality program production. We are doubtful, however, that adjustment for all these factors would bring the average ETV rating up to the critical minimum we estimate for marginally

profitable network programming. Our doubts arise in part from the observation that those network programs which come closest in content to what seems to be desired by minority audiences rarely obtain ratings in excess of ten on their first network showing. However, some network nonentertainment specials have obtained ratings in the 10-15 range.

Currently, networks typically delete programs which fail to achieve a wintertime prime-time rating of at least 15. A favorable assumption to the prospects for increasing diversity through expansion is that total viewing audience would not be increased with additional broadcasting of presently popular program types. With this assumption, together with the unrealistic assumption that revenue per viewing household would not decrease with increasing competition, the indication is that adding one or two networks would increase the number of programs with ratings of 10 to 15 which would be broadcast. Most regular entertainment programs which fail to survive in the present three network structure earn ratings in this range. It is thus highly likely that such an expanded system would continue to cater overwhelmingly to audiences in search of light entertainment programming. Furthermore, while a fourth nationwide network is probably technically feasible, its economic feasibility is questionable. And a fifth is technically infeasible, given existing channel assignments.

As an alternative, high-culture programming might be provided on a syndicated basis to affiliates and independent stations. Ignoring distribution costs, the above calculations for network showings of this type of programming indicate that a producer of a first-run syndicated program which was expected to reach 9 million TV homes would have to price the program so that he earned \$9.50 per thousand viewers in order to recover production costs. But it can be shown, assuming the minimum rating for regular network programming is 15, that an affiliate would not delete even the least attractive network program in favor of a syndicated program with an expected rating of 10 unless the effective price of the syndicated program was less than \$6.00 per thousand.

This means that a syndicator of such programming would have to rely on sales to independent stations for showings in prime time. If we were to assume an independent station in every market, and no UHF disadvantage, total viewing on a syndicated basis would be the same as on a network basis, provided the program got two showings in each market.

First-run syndications presently attract an audience of 5 percent of TV homes on a VHF station in a four-station market and their average cost is about \$3.50 per thousand viewing homes per half-hour episode. Using the estimate of the value of viewing homes to independent stations which Park presents in the RAND report which his paper summarizes, we estimate an independent would earn about 1.6 times as much net revenue (after program costs) by showing the high-quality, minority-audience program.

This seems to indicate that having four stations in all markets, with no UHF disadvantage, would make a substantial contribution to diversity. Unfortunately, that is not likely to be the case. For such an expansion in the potential audience of unaffiliated stations would markedly change the economics of production for the first-run syndication market. Specifically it would become profitable for program packagers to produce, for first-run syndication, programs of the type which are now deleted by the networks after one or one-half season. These programs typically do not now find their way into the syndication market because as off-network programs too few episodes have been produced to make them an attractive buy for independents and their potential market is at present too small to make it profitable to produce additional episodes for syndication. However, further production would become profitable with the expansion of the potential market of independent stations assumed here. As a result it seems most likely that conventional programming would dominate the schedules of independent stations in such an expanded system.

Cable television is a means of escaping the technological constraints and the popular belief is that growth of cable television will tend to increase diversity. But, for reasons cited by Park, FCC policy has impeded cable

growth by restraining distant signal importation in the larger markets. Park's imaginative and valuable study finds this policy to be misdirected because the impact of cable penetration would be relatively minor in the markets which present policy protects and relatively severe in the small markets which present policy does not protect. Although some of his estimation procedures raise questions in our minds, we shall refrain from discussion of these since his larger study has not been made generally available to this audience and because we doubt that the substance of his major conclusion would be much altered by taking account of our reservations.

An important implication of Park's study is that it may well be difficult to devise a policy which will preserve local station service to viewers in small markets while at the same time providing a diverse menu of viewing options. At the same time his study leaves unanswered the question of what kind of diversity would emerge if CATV's were to achieve ultimate penetration in the larger markets. This is not meant as criticism of Park's work, but only to point out that much remains to be done.

Discussions of broadcast performance have in our view been overly dominated by the implicit assumption that technological constraints on channel expansion arising from spectrum scarcity and the inferiority of UHF signals constituted the overriding barrier to improvement. As a result, means of overcoming those technological constraints, such as CATV, tend to be viewed as panacea. And yet, broadcasters' behavior and performance cannot be explained and policy alternatives cannot be evaluated without consideration of the economics of program production and supply. While the immediate impact of relaxing technical constraints on broadcast capacity is to reduce the opportunity cost of broadcast time, that effect feeds back on the economics of program production and supply, and this full chain of events must be considered in assessing the impact of expansion in channels on performance. Thus, it is misleading to estimate that impact from regression analysis of variations in diversity in present markets, no matter how sophisticated the measure of di-

CONGLOMERATE AND VERTICAL RESPONSES TO MARKET IMPERFECTION

Expectations, Mergers and Equilibrium in Purely Competitive Securities Markets

By JOHN LINTNER
Harvard University

The first part of this paper summarizes and contrasts the differing principal sources of investors' gains in each of our major merger movements. The following more theoretical section examines the potential gains in "pure" conglomerate mergers characteristic of the last decade. The *criterion* of gain is an increase in the equilibrium aggregate market values of the combining companies at the time of and due to the act of merger *ceteris paribus*, exclusive of promoters' "cuts."

I

Mergers Before 1920. The two great waves of consolidation before World War I, broken only by the depression of 1893 and Supreme Court action on the Trust device, can appropriately be considered together. These much studied¹ massive movements clearly created much of our twentieth century pattern of corporate concentration. In the more noteworthy, larger, and important (if not in the majority²) of individual cases, there can be equally little doubt that "monopolization" was the dominant *business* motive for the merger. Although outside "promoters" seem to have played a negligible role in

some of the larger and more successful mergers [19, p. 163], it is clear that the vigorous activity of promoters and underwriters interested in their profits provided a major stimulus to many of the combinations in this period.³ "Rationalization," technological change, and the anticipation of added profits from economies of scale have also been frequently cited as additional important factors, but recent research concludes that their role was probably rather secondary in the overall movement [24, pp. 71-106] and [18, pp. 46-69]. The emergence of large-scale and well-organized securities markets was an essential precondition for the large security issues involved in many of these early mergers [27, p. 28]—and the rapidly increasing supplies of securities for trading produced by these consolidations significantly accelerated the development of the securities markets themselves in a mutually reinforcing sequence [24, pp. 89-96]. Regarding mergers through World War I, Nelson further concludes that "the high correlation between merger activity and stock prices suggests that much of the merger activity of the period had its origin in, or was influenced by, the stock market. Further examination indicated that capital market factors overrode the level of in-

¹ For summaries, see [3, Chap. X], [19], [24], [18, Chap. II] and consult basic references cited in these.

² See Nelson [32, pp. 101-3]. But see also Markham [24, pp. 154 ff.] and Geo. W. Stocking's comments at pp. 191 ff., together with the primary sources cited.

³ See [11], [3, pp. 310 ff.], [18, pp. 38-41] and underlying references.

dustrial activity in influencing merger activity" [24, p. 105].

In most, if not all, of these mergers, the aggregate market value of the securities of the new consolidation substantially exceeded the sum of the values just before the merger of the securities of the firms being merged. Even the most simplified versions of modern capital market theory would, of course, predict such step-ups in values through merger under the stated conditions. If we let the aggregate value V_i of all the securities of the i th firm outstanding just before the merger be a simple capitalization of its expected earnings before interest $V_i = \bar{X}_i / \rho_i$, and the total value of the merged firm be $V_M = \bar{X}_M / \rho_M$, we may consider the simplest case in which all constituent firms and the consolidated enterprise are in the same "risk class"⁴ so that $\rho_i = \rho_M = \rho$. Let the aggregate profits of the merged firm be $100h\%$ greater than the sum of the profits of the combining units due to the (product) market power created by the consolidation, and also allow for the added operating profits of $100s\%$ expected in some mergers due to cost savings by writing $\bar{X}_M = (1+h)(1+s)\Sigma_i \bar{X}_i$. The aggregate market value of the securities of the combine

$$(1) V_M = (1+h)(1+s)\Sigma_i V_i > \Sigma_i V_i$$

since $h \gg 0$ and $s \geq 0$.

Although securities of the new merger were often said to be "watered" to the extent of this step-up in aggregate values,⁵ the capitalization of monopoly profits alone would leave capital gains in the postmerger equilibrium for owners of the constituent companies even after a 100 $b\%$ "cut" to the promoter-midwives so

⁴ By holding risk class invariant across any merger, we eliminate any consideration of "risk sharing" gains which are analyzed separately in the final section. We also simplify the analysis here by reserving tax gains from debt interest to the final section.

⁵ See [4], [5], and [26], but note [25, p. 113].

long as $h > b / (1-b)$. Previous security holders would have been left in an improved financial position in terms of *current* market values whenever *they* (independently and collectively) *assessed* as much as a 25% increment in total ex-merger *expected* profits due to enhanced monopoly power and *expected* cost savings from merger, even when the "financial entrepreneurs" absorbed a 15%–20% share of the successor company, as often happened.⁶ This conclusion is strictly independent of any further increase in "early-on" market value attributable to manipulation, "touting," padding promoter's projections, or overselling—and it is also, strictly speaking, independent of the subsequent "performance" of the merged companies in providing greater returns to investors. Securities markets equilibrate on the basis of the *ex ante* judgments of the time. Information becoming available later regarding actual operating results affects the investment performance of securities over time *after* the mergers. We know that along with many well-known successes, there was a rather high incidence of subsequent failure in these mergers.⁷ If comprehensive studies of subsequent investment performance would show subnormal risk-adjusted returns relative to the market, we would have to attribute the deficiencies to (perhaps systematic) biases at the time of the mergers in investors' *ex ante* assessments of necessarily uncertain *ex post* re-

⁶ See [11], Chap. XII, [4], and [26, p. 64].

⁷ Dewing [4] studied 35 major early consolidations and generalized that "the trusts turned out ill" [5, p. 734]. Livermore [16] found that about 45% of each of two much larger groups of mergers with and without major market power were "successes" through 1932, but the rest "limped" or failed. Nelson [24, pp. 96–9] found that 13 large mergers of 1899–1901 produced ten year investor returns 2.1%–3.6% greater than bond yields at the time, but (though Nelson did not make the comparison) they were below average returns on all common stocks (without adjusting for risk).

sults, or to "overtouting" and manipulation, or to some combination of the two sources. But all such results of possible further studies would be consistent with our conclusion of general (or even large) gains in values at the time of the mergers.

The strong correlation of merger activity with stock market levels can also be readily explained in terms of our simple cross-sectional equilibrium model by the widely noted dependence of expectations of the future upon recent experience and the greater ability of rising markets to absorb large new issues. In particular, even under strict random walk or martingale assumptions [6], higher levels of general stock market prices imply either that expected future profits have been adjusted upward or that capitalization factors have been reduced by lower interest rates or smaller *ex ante* risk assessments. Enhanced estimates or expectations of "em-merger" magnitudes imply greater (absolute) gains from the act of merger given unchanged estimates of "*k*" and "*s*". If judgments of the degree of monopoly power and relative cost savings themselves are also cyclically variable, the *relative* market value gains anticipated from merger would also be greater in booming than in depressed stock markets.

Mergers in the 1920's. With prosperity and rising stock markets, mergers steadily increased during the 1920's to a peak of 1,245 disappearances in 1929—a pace matching the earlier watershed year of 1899. Stigler characterized the mergers of the 1920's as "meggers for oligopoly" in contrast to the earlier mergers for monopoly when "the leading firms seldom merged less than 50% of the industry's output; in the later period the percentage has hardly ever risen that high" [27, p. 31]. At the same time, the record is clear that this "era of consolidation" substantially tightened the oligopolies and increased concentration in such major in-

dustries as steel, oil, machinery, and copper [3, pp. 297–9], and that "in many fields mergers reduced to a mere handful the number of sellers whose policies really mattered in shaping the market" [28, p. 41]. For present purposes, we simply note that even if Markham's conclusion [19, p. 169] that "the larger horizontal mergers (of the 1920's) on balance may well have stimulated as much competition as they stifled" is accepted without qualification,⁸ the *anticipation* of added profit through improved market power and competitive position surely was a very substantial factor in many of the mergers in this period, though probably rather smaller than in the massive consolidations earlier. The role of external promoters also seems to have been more modest than at the turn of the century, but they continued to be a significant force during the 1920's [5, p. 929], [29, pp. 85–6].

In addition to the marked shift from single multi-firm consolidation to "piecemeal absorption" of rather smaller firms in most cases, there was a very marked increase in "quasi-conglomerate" mergers among firms selling the same product in noncompeting territories (e.g., food or bakery chains), or producing substantially different lines of products (although generally with complementarities in production or distribution, or some significant cross-elasticity in product demand).⁹ In the twenties, there was also more emphasis than earlier on economies of large-scale distribution, advertising, and sales promotion, as well as economies in production, while improving technologies in transportation and communication led to some mergers by greatly increasing the size of markets.

⁸ Compare Markham [19, pp. 167–73] and Stocking's following "comments" (pp. 204–11). Also see Tippetts and Livermore [30, p. 373] and Bain cited in [3, p. 307].

⁹ These quasi-conglomerate mergers accounted for over 25% of the total [8, p. 63].

Apart from perhaps a few "pure" conglomerate acquisitions in the modern sense, the other significant differences in the merger movement of the 1920's fit rather easily into the very simple theoretical model introduced above. The evidence suggests that the percentage gain in expected profit due to incremental "monopoly" or "oligopoly" power (our "*h*" factor) was probably generally somewhat smaller than was common at the turn of the century, and it may have been negligible in more cases than in the earlier movement. In a significant number of cases, however, this market power factor was enhanced by the ability to influence (or control) prices of different products with significant cross-elasticities of demand [25]. Added profits from market expansion *per se* may be included along with those from economies of scale in production, distribution, and promotion in our "*s*" factor, and these together appear to have been substantial *ex ante* in a larger fraction of cases than before. Promoters' shares "*b*" were small or zero in many mergers and much smaller than in the earlier movements in most of the other cases. Our criterion for the equilibrium value of the securities (held right after the merger by the owners of the predecessor companies) to be greater than their equilibrium value *ceteris paribus* without the merger is now:

$$(2) \quad h + s + hs > b/(1-b). \\ s > 0; h, b \geq 0.$$

Mergers in 1940-47. Merger activity picked up markedly during and after World War II, with over 2,000 firms with \$3.5 billion in assets acquired in reported mergers. Serious concerns were expressed regarding the effects on concentration [7, pp. 25, 28, 68-9], but in contrast to earlier patterns, the *relative* growth of larger acquiring firms by merger was systematically and markedly less than the growth of smaller firms, and apart from

limited exceptions in a few industries, increases in absolute concentration were minor and relative concentration was actually reduced.¹⁰ In terms of our primary interest here in securities markets and mergers, we particularly note that, although stock prices were rising strongly to 1946 as merger activity increased, there was little outside financing involved in the mergers in these years and the usual *causal* relation with stock market levels was essentially *reversed*. One of the important reasons for acquisition rather than internal expansion was the fact that the stocks of so many companies were selling substantially *below* their book values and even more below their reproduction costs. It was therefore often cheaper to acquire the desired facilities, products, and market outlets by buying another company rather than building or developing them directly. (It was also, of course, quicker and avoided intensifying competition as direct entry would have done.) The fact that merger activity did not drop off with the market break of 1946 is probably explained¹¹ by the further fact that sellers were taking the initiative far more frequently than buyers in this period [3, pp. 14-20].

In short, aggregate market values were clearly enhanced by the mergers in this period. In terms of our earlier criteria, the "*s*" term was usually large, though it generally took the form of (a) bargain purchases below reproduction costs of assets acquired, and (b) opportunity gains in the avoidance of the costs, risks, and de-

¹⁰ See [3, pp. 241-80] and further discussions in *Review of Economics and Statistics*, Feb. 1951, as well as [19, pp. 174-80] with Adams' comments following on pp. 182-90.

¹¹ Nelson in Alberts and Segall, *The Corporate Merger* (see [1] above), finds a much weaker correlation of merger and the stock market in 1919-61 than he had found before 1920 in [24]; but the aberrant and "outlying" observations of 1941-8 explain most of the reduced covariation over the whole period (cf. pp. 57 and 63).

lays which would have been involved in the alternative course of developing a new plant, product, organization, or market. Although the "*h*" factor was rather large in a few cases, it was generally small or moderate in size and usually took the form of avoiding the more intense competition which would have been created by the buyer's direct entry into the market, rather than the affirmative act to create additional (product) market power through merger. With $h > 0$, $s > 0$ and b negligible, the aggregate value gain is obvious. Sellers clearly participated in the overall gain because of a selling price higher than their alternatives afforded, as well as tax benefits in an important minority of cases.

Mergers since 1950. After a lull with weaker markets in 1948–9, merger activity expanded to the 1925 pace by 1955. After a small dip again in 1957–8, they moved to a new higher plateau in the early 1960's (just below the 1928–9 pace), before virtually exploding with the booming stock markets of 1966–68. In each of the latter two years, the fraction of all manufacturing and mining assets absorbed by merger exceeded the fraction in 1929. As in earlier movements, larger firms were relatively more active acquirers throughout; but in contrast to the 1940's when the acquisitions of larger firms were relatively small, there has since been a very rapid increase in the number of relatively large firms being acquired, and *overall* concentration has clearly been increasing substantially due to mergers in recent years.¹²

One of the distinctive features of this longest and most massive merger movement has been the rapid increase in "conglomerate" forms of consolidation to include over 80% of all merger assets in 1964–8 [8, p. 673]. But 64% of all assets

in conglomerate mergers involved "product-extensions" of companies "functionally related in production and/or distribution" or "geographic market extensions" (e.g., bakery chains)—and *these* types of conglomeration have very substantial historical precedent [8, pp. 60, 63, 673]. Including the standard "horizontal" and "vertical" categories, these more traditional types of merger still accounted for virtually 80% of all acquired assets through 1967. Preoccupation with the (relatively) newer and more exotic "other" conglomerates—i.e., those with no buyer-seller relationship nor a functional relationship in manufacturing or distribution—seems to have obscured the tremendous scale of more traditional types of merger activity over the last decade or so. *Exclusive of "other" conglomerations*, more than \$10 billion assets were acquired in 1952–59, and another \$29.5 billion in the nine years 1960–68; taken together, these involved about *the same percentage of all manufacturing assets as characterized the roaring 1920's*.¹³

Nevertheless, the great increase in "other conglomerate" mergers from small beginnings to a crest of \$5.5 billion in 1968 is the most distinctive feature of this otherwise large and massive movement. But even though the F.T.C.'s "other conglomerate" category excludes "product and market extensions" by definition, some of these mergers may also have opened up significant opportunities to enhance preexisting positions of discretionary power in product markets (by reciprocity, exploiting cross elasticities of demand, etc.). Any such features of these "other conglomerate" mergers may raise impor-

¹² This paragraph is based on [8, esp. pp. 38–54, 161–84], and see also [25, pp. 108–11].

¹³ Data on mergers by type in 1952 ff. from [8, p. 673]; this \$39.5 billion was 76.3% of all assets acquired. Since all mergers summed to 21.1% of total manufacturing and mining assets, 16.1% may be ascribed to more traditional types of merger since 1952. The corresponding "total disappearance" in 1921–31 was 16.3% [8, p. 666].

tant issues of public policy and Anti-Trust enforcement, but such sources of gain in security values fit easily into our "*h*" factor and raise no new analytical issues for this paper. Similarly, to the extent that *ex ante* improvements in unmerged profits are anticipated from cross-breeding technologies, or from introducing more effective and modern systems of internal control or "better management," the resulting gains from merger are on a par with anticipations of other operating economies (our "*s*" factor) and raise no new problems for a theory of security market values or for "finance." There were also more surely "*h*" and "*s*" elements of *ex ante* significance involved in many of the more traditional types of merger covered in the previous paragraph—and all such elements in any merger, however classified, will lead to an increase in aggregate market value at the time of the merger by standard analysis. In this context, the most important analytical issues raised by recent merger activity consequently focus on the question whether the aggregate market values of firms can be raised by merging them, even when there are no gains in combined operating profits anticipated from increments of (product) market power, economies of scale, or other improvements in efficiency—i.e., when both our "*h*" and "*s*" factors are zero.

For purposes of the usual analyses of Industrial Organization, it would be sufficient to define a "pure" conglomerate merger as one satisfying the condition $h=s=0$ in some objective or *ex post* sense. But since security values depend upon assessments of an uncertain future, we must refine this definition to require that the *expected* value \bar{X}_M of the net operating profits (before interest and taxes) of the merged firm be identically equal to the sum of the expected net operating profits of the separate firms being merged. In addition, to allow for anticipated growth and other

changes in conditions, and since different investors will have different *ex ante* assessments, the condition must hold for all future years for all *i*th investors holding securities in each of the firms being merged. Our basic definition of a "pure" conglomerate merger thus requires that $\bar{X}_{Mitk'} = \sum_i \bar{X}_{itk'}$ for all investors *k'* and all $t \geq 0$. Finally, for some purposes when different subsets of investors hold securities in the firms being merged, we must supplement our basic definition with the added condition that *every* investor's assessment of all \bar{X}_{Mit} be made on the condition that $h=s=0$.

Given that these conditions are satisfied, we may note five distinguishable sources of increase in the aggregate market value of (all the securities of) the merging firms at the time of and due to the act of merger *ceteris paribus*, when securities markets are in purely competitive equilibrium just before and after the merger. (Otherwise, "bargains" would be a sixth source.)¹⁴ In particular, we identify: (1) Gains from favorable tax treatments incident to certain mergers *per se*; (2) Gains from greater leverage, and/or lower borrowing costs due to size; (3) Possible gains from merging imperfectly correlated income streams to preserve expected returns with reduced "risk"; (4) Enhanced opportunities through mergers to manipulate or legally exploit the "aura of discretion" within "generally accepted accounting principles" in reports to shareholders; and *independently* (5) Automatic and instantaneous gains in earnings per share in an important class of mergers.

"Dirty Pooling," suppression of asset costs at the time of merger to pad subsequent earnings, and other accounting devices which are likely to mislead many shareholders, clearly raise issues of the

¹⁴ For discussion see [13], [10], [1], and [2].

greatest practical importance, but these are not our concern here. For present purposes we will simply *assume* the widespread dissemination by all firms of reasonably complete accounting reports characterized by "full disclosure," freedom from any "mirage created by accounting,"¹⁵ and meeting high standards of objectivity.

Taxes. The Federal Trade Commission finds "Tax exemption of corporate reorganizations has the most pervasive impact of all tax provisions favoring mergers" [13, p. 143]. This "tax free" treatment, however, involves inequity only *vis-a-vis* alternative (taxable) sales, and is not a positive incentive to merger *relative to* continued independent operation. Mergers *are*, of course, stimulated (and combined market values raised) under present tax law whenever either firm has tax loss carry forwards or investment credits with no comparably large and otherwise taxable income in sight as an independent unit. But the use of credits or offsets against later taxable income in place of cash refunds is regarded as an imperfection on *general* grounds in the Theory of Public Finance,¹⁶ and the stimulus to merger would be absent in an otherwise neutral tax law. Other provisions permit the deferral of taxes on limited kinds of income in certain mergers, and the present value of the deferment will also raise security values. To focus the rest of the analysis, we will simply assume henceforth that tax considerations are neutral except for tax savings on any corporate debt incident to mergers.

Leverage. Apart from "debt-equity switching" to manipulate income statements [8, pp. 149-50], there are two sources of gain through leverage from

merger *per se*. (We assume that the merging firms already have optimal (entity-value maximizing) mixtures of debt and equity in their capital structures.)¹⁷ The first arises from the fact that borrowing costs decline with size of firm, other things equal, even in idealized markets under uncertainty where information itself is an economic good—because of "lot size" scale economies in credit investigations and security issue costs as well as "marketability."¹⁸ Large firms can thus refinance debt of small independent firms at lower economic cost resulting in a genuine capital gain through merger. An additional and independent¹⁹ source of gain through lower debt costs has very recently been pointed out by Levy and Sarnat [12] and developed (independently) by Lewellen [13]. As the former explain,

"If we assume that in any given year (or run of years) there exists for each individual firm some positive probability of suffering losses large enough to induce financial failure, . . . the joint probability of such an event is reduced by . . . the combination of other than perfectly correlated income streams in a conglomerate merger. . . . The diversification can be expected to create a true economic gain to the shareholders owing to the fact that the combination of the financial resources of the two firms making up the merger reduces lenders' risk while combining each of the individual shares of the two companies in investors' portfolios does not." [12, p. 801].

Gains from Merger in All-Equity Case. In order to simplify the rest of the analysis, we henceforth assume all firms have all-equity capitalizations. It is clear from Markowitz's portfolio theory that all risk-

¹⁷ See [8, pp. 134-8] and references in fn. 14.

¹⁸ See [8], S.E.C. Reports on Flotation Costs, and Federal Reserve Bulletins for studies of costs by size of loan.

¹⁹ This source of gain involves an increase in "debt capacity" (more debt at same risk), while the previous point lowered the cost of a given (total) volume of debt.

¹⁵ The phrase is the Trade Commission's [8] at p. 129; for a discussion of some of the major issues involved here see pp. 120-32 and references there cited.

¹⁶ See Musgrave's *Theory of Public Finance* or any comparable reference.

averse investors will prefer a mixture of two imperfectly correlated income streams to either stream separately, and company mergers produce such a mixture. Alberts [1], [2], however, has argued and Myers [23] and Levy and Sarnat [12] have later proved that in *perfect* capital markets the value of the merged firm will be the same as the combined market values of the separate firms—because portfolio diversification by investors will have already exhausted the diversification benefits available through corporate merger. This conclusion is strictly valid *provided* “perfect” capital markets are defined to require that *every* investor invests in *every* stock available in the market. Indeed, this particular conclusion regarding mergers holds even more generally,²⁰ so long as identical subsets of investors are assumed to hold the stock of *each* of the merging firms. Otherwise, the conclusion of “no gain” from merger need not (and very generally *will not*) follow.

The essential reason is that, as shown in [14, pp. 390 ff.], the market price of risk will be the same for each security when purely competitive markets are in equilibrium *if but only if* every investor has *some* (long or short) holding of every security in his (optimized) portfolio. But the New York Stock Exchange Shareholders’ Surveys show that investors have an average of less than four stocks in their portfolios (and mutual funds and other institutional investors each hold only a small fraction of outstanding issues in any one portfolio). And under conditions where

different securities are held by different subsets of investors, the market price of risk for any *i*th stock varies *inversely* with the summation of the risk-tolerances (reciprocals or risk-aversion coefficients) *of the investors who have it in their portfolios* (long or short). In general, when there is a riskless asset yielding a rate of return r^* , the aggregate market value of any security in equilibrium may be written.

$$(3) \quad V_{oi} = n_i P_{oi} = (n_i \bar{P}_{ii}^a - \gamma_i R_i^a) / (1 + r^*),$$

$$\gamma_i = H_i^a / N_i,$$

where n_i is the number of shares of the *i*th stock outstanding; \bar{P}_{ii}^a is the relevant weighted average of the assessments of expected end of period price per share of the investors who hold the stock; R_i^a is n_i times the relevant weighted sum of the *marginal* contributions of the *i*th stock to the *portfolio* variance of the respective individual investors who hold the stock (long or short);²¹ γ_i is the “market price of risk relevant to the *i*th stock,” which equals the harmonic average H_i^a of the risk-aversion coefficients of the subset of investors who hold the stock divided by the number of such investors N_i .

Consider two firms whose investors collectively assess the *same* expected end-of-period aggregate value

$$(\bar{V}_{ii} = n_i \bar{P}_{ii}^a = \bar{V}_{ij} = n_j \bar{P}_{ij}^a),$$

and let the collective assessment of aggregate portfolio risk attributable to the two securities be the same ($R_i^a = R_j^a$). Also let the harmonic means of the respective subsets of investors be the same ($H_i^a = H_j^a$). Nevertheless, *if firm i has more shareholders than j, it will have more value in the market*—because, *ceteris paribus*

²⁰ Myers covers the more general case in a “state-preference” model, where the critical assumption is that there exists equivalent securities at the margin for all of a given set of investors. The condition stated in the text is also critical in Myers’ proof (see second sentence on p. 11 in his paper). Although we are only concerned with mergers here, we note explicitly that the ensuing arguments in the text apply equally to capital budgets, showing it is *not* a matter of indifference which firm undertakes a given project.

²¹ The “R” term therefore includes *all covariances* within all portfolios of investors who hold the *i*th stock.

$N_i > N_j \rightarrow \gamma_i < \gamma_j \rightarrow V_{oi} > V_{oj}$.²² In particular, if i has one million shareholders and j only has ten thousand shareholders but the above *ceteris paribus* conditions hold, then V_{oi} will be larger than V_{oj} by an amount equal to 99% of the present value of the aggregate "risk discount" of the j th firm (i.e., $V_{oi} - V_{oj} = .99\gamma_j R_j^a$ when $R_i^a = R_j^a$, $H_i^a = H_j^a$ and $\bar{V}_{ii} = \bar{V}_{ij}$).

Now consider the merger of two companies, using for simplicity the *ceteris paribus* equalities of the previous paragraph. (1) If in addition, $N_i = N_j$ and the membership of the two shareholding groups is identical, then there would be no change in market value from the merger. (2) Even if $N_i = N_j$, but membership is not identical and the merged shareholder list $N_{(ij)} > N_i$ or N_j separately, then the price of risk for the merged company $\gamma_{(ij)} < \gamma_i$ or γ_j , and aggregate market value of the merged firm will be greater than the sum of the separate values of the merging companies, $V_{o(ij)} > V_{oi} + V_{oj}$, on the very reasonable added assumption that the (weighted average) assessments by i shareholders of j 's independent prospects—which i 's owners use in assessing the prospects of the merged firm (ij) after the merger is announced—not be inferior (at the new price of risk γ_{ij}) to those of j 's own shareholders, and *vice versa*. Finally, if $N_i > N_j$, the same conclusion follows even more directly.

To this point we have assumed that the " j " company's shareholders who did not hold any " j " stock before the merger (and " j " holders not owning any " i " stock) simply had no assessments of the "other" company's prospects before the merger was announced. It is surely a fact that

²² The rationale of the result is that the market price of risk varies inversely with the (relevant or effective) size of market because of the efficiency of purely competitive securities markets as risk-eliminating as well as risk-sharing mechanisms. See [15], and [14, pp. 393–98].

most investors at any one time can assess a probability distribution over only a few stocks—even the largest institutions have "considered judgments" on no more than 100–200 companies at a time (out of 1.5 million corporations filing tax returns). If "perfect" securities markets by definition require conditions which insure that all investors have some positive (or negative) holding of all securities (and otherwise the "market price of risk" is not the same for all securities [14]), then we have merely shown that a certain very prevalent "imperfection" does produce genuine gains in economic values from "pure" mergers. But the only "imperfection" required for this result is unavoidable, since it is the (consequence of) regarding information as an economic good, recognizing that getting information involves economic costs, and that there is an opportunity-value to the time devoted to security appraisal. If such markets are not perfect, they are surely the most nearly perfect attainable under uncertainty.

One final observation on this case: if (contrary to previous assumptions) the difference in the shareholding groups before the merger arises after thorough investigations of both companies by both groups of investors, and the non-holders are simply "bears" on the respective stock who cannot sell short, then a "pure" merger may reduce (rather than increase) market values—but it is unlikely as a practical matter to do so.²³

²³ The market value of any stock in equilibrium depends only on the assessments of those who have it (long or short) in their portfolios [14, p. 394], and the assessments of different investors are in effect weighted by the fraction of the total supply of the security they hold [14, pp. 359–60, 396–7]—which incidentally points up the common error of arguing in terms of mythical "marginal" investor (as, for instance in [22]). The particular conclusion in the text here follows from observing (a) that the total short sales of the "frustrated bears" on the stock are unlikely to have been any large fraction of the total outstanding, and (b)

Automatic and Instantaneous Increases in Earnings per Share from P/E Ratio Differences. Even in the "all equity capitalization" case with no increase in combined total earnings, and with *no* manipulations of accounting,²⁴ there is an automatic increase in the earnings *per share* of the acquiring company whenever its price-earnings ratio is higher than that of the acquired firm. Expectations of the future generally reflect more or less recent past experience, and many investors judge growth and growth prospects in terms of earnings per share. The merger as such thus tends to raise the assessment of future share earnings (and hence the expectation of future share prices) in the minds of many investors. Since judgments of appropriate current share values are derived from assessments of future share values, which are higher with than without the merger, the current aggregate value of the merging companies in equilibrium will be raised, *ceteris paribus*, even in a "pure" conglomerate case in *strictly perfect* securities markets with the market price of risk the *same* for *all* securities.²⁵ More-

any "down draft" from including their assessments after the merger is very likely in any case to be more than offset by the "updraft" developed earlier.

²⁴ The F.T.C. also distinguishes these "instantaneous gains in earnings per share" (pp. 133-4) from "asset cost suppression" and other instances of "accounting as a tool of deception" [8, pp. 121-32].

²⁵ Although arguments have been made that market values "should" not be increased in this case (e.g., [21, p. 23]), the result is not only widely observed in practice, but, *given* the way investors seem to form their *ex ante* assessments, it also follows in strict theory. Wise men may know there is no perpetual motion and that "chain letters" come to an end, affecting *later* investment performance. But not all men are wise and the market uses the concurrent *ex ante* assessments of *all* men (holding the stock), made in whatever way *they* make them, to determine its prices right after the merger. Keynes' classic description of the stock market is still as relevant as when he wrote it. (In this respect, the 44% per annum "performance" of 28 "acquisitive acquirers" in Lynch [18, Chap. III] and the results of Lorie's [17] study of "funny money securities," both through 1967, may be compared with the 1969-70 "action"!)

over, since companies with more rapidly growing earnings per share tend to have "high- β " stocks, the dispersion of price-earnings ratios increases with the level of the general market, with the result that this type of merger also contributes strongly to the frequently observed correlation of merger activity with levels of stock market prices.

Summary. Each of our major-merger movements has had a different characteristic pattern of primary sources of investor gains. Direct gains from "monopolization" have generally diminished in relative importance while anticipated economies of scale and improvements in efficiency have varied in relative importance from movement to movement (and within each separately). Even "pure" conglomerate leads to investor gains in *perfect* capital markets due to (a) reductions in lenders' risks of bankruptcy losses, (b) scale diseconomies in credit investigations of smaller firms, in flotation expenses in public issues, and in investor information ("marketability"), and (c) changes in investors' assessments of future prospects per share when P/E ratios differ. Moreover, there are also gains in all-equity cases of "pure" company diversification when different subsets of investors hold the stocks of the merging companies and security markets are otherwise perfect. All the gains listed are investor gains at the time of and due to merger, other things equal (quite apart from any effects of "touting" or "manipulation").

REFERENCES

1. Albers, William W., "The Profitability of Growth by Merger" in W. W. Albers and Joel E. Segall, *The Corporate Merger*, Chicago Univ. Press, 1966.
2. Albers, William W., "The Profitability of Conglomerate Investment Mergers: Sources and Prospects," *Univ. of Washington Business Review*, Winter 1970.
3. Butters, J. Keith, John Lintner and Wm.

- L. Cary, *Effects of Taxation on Corporate Mergers*, (Division of Research: Graduate School of Business Administration, Harvard University, Boston, 1951).
4. Dewing, A. S., "A Statistical Test of the Success of Consolidations," *Quarterly Journal of Economics*, November, 1921.
 5. Dewing, A. S., *Financial Policy of Corporations*, Ronald Press, 5th ed., 1953.
 6. Fama, Eugene, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, May, 1970.
 7. Federal Trade Commission, *The Merger Movement, A Summary Report*, U.S.G.P.O., 1948.
 8. Federal Trade Commission, *Economic Report on Corporate Mergers*, Hearings, Subcommittee On Antitrust and Monopoly, Committee of the Judiciary, U.S. Senate, 91st Congress, 1st session, Part 8A, U.S.G.P.P., 1969.
 9. Fisher, Lawrence, "Determinants of Risk Premiums on Corporate Bonds," *Journal of Political Economy*, June 1959.
 10. Gort, Michael, "An Economic Disturbance Theory of Mergers," *Quarterly Journal of Economics*, November 1969.
 11. Jones, Eliot, *The Trust Problem in the United States*, Macmillan, 1929.
 12. Levy, Heim and Marshall Sarnat, "Diversification, Portfolio Analysis and the Uneasy Case for Conglomerate Mergers," *Journal of Finance*, September 1970.
 13. Lewellen, Wilbur G., "A Pure Financial Rationale for the Conglomerate Merger," forthcoming in *Journal of Finance*, May 1971.
 14. Lintner, John, "The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Securities Markets," *Journal of Financial and Quantitative Analysis*, December 1969.
 15. Lintner, John, "The Market Price of Risk, Size of Market, and Investor's Risk Aversion," *Review of Economics and Statistics*, February 1970.
 16. Livermore, Shaw, "The Success of Industrial Mergers," *Quarterly Journal of Economics*, November 1935.
 17. Lorie, James H. and Paul Halpern, "Conglomerates: The Rhetoric and the Evidence," *Journal of Law and Economics*, April 1970.
 18. Lynch, Harry H., *Acquisitive Conglomerate Performance*, unpublished thesis, Grad. School of Bus. Adm., Harvard Univ., 1969.
 19. Markham, Jesse W., "Survey of the Evidence and Findings on Mergers" in *Business Concentration and Price Policy* (N.B.E.R., Princeton Univ. Press, 1955).
 20. Markham, Jesse W., "Antitrust and the Conglomerate: A Policy in Search of a Theory," *St. John's Univ. Law Review*.
 21. Mead, W. J., "Instantaneous Merger Profit as a Conglomerate Merger Motive," *Western Economic Journal*, December 1969.
 22. Mueller, Dennis C., "A Theory of Conglomerate Mergers," *Quarterly Journal of Economics*, November 1969.
 23. Myers, Stewart C., "Procedures for Capital Budgeting Under Uncertainty," *Industrial Management Review* (M.I.T.), Spring 1968.
 24. Nelson, Ralph L., *Merger Movements in American Industry, 1895-1956* (N.B.E.R., Princeton Univ. Press, 1959).
 25. Scherer, F. M., *Industrial Market Structure and Economic Performance*, (Chicago, Rand McNally, 1970).
 26. Seager, H. R. and C. A. Gulick, Jr., *Trust and Corporation Problems* (New York, Harper and Brothers, 1929).
 27. Stigler, George, "Monopoly and Oligopoly by Merger," *American Economic Review*, May, 1950, pp. 23-34.
 28. Stocking, Geo. W. and Myron W. Watkins, *Cartels for Competition?* (20th C. Fund, 1948).
 29. Thorp, Willard L., "The Persistence of the Merger Movement," *American Economic Review*, Suppl., March, 1931.
 30. Tibbetts, Chas. S. and Shaw Livermore, *Business Organization and Control*, Van Nostrand, 1932.
 31. Weston, J. Frederick, *The Role of Mergers in the Growth of Large Firms* (Univ. of Calif. Press, 1953).

The Vertical Integration of Production: Market Failure Considerations*

By OLIVER E. WILLIAMSON
University of Pennsylvania

The study of vertical integration has presented difficulties at both theoretical and policy levels of analysis. That vertical integration has never enjoyed a secure place in value theory is attributable to the fact that, under conventional assumptions, it is an anomaly: if the costs of operating competitive markets are zero, "as is usually assumed in our theoretical analysis" (Arrow, 1969, p. 48), why integrate?

Policy interest in vertical integration has been concerned mainly with the possibility that integration can be used strategically to achieve anticompetitive effects. In the absence of a more substantial theoretical foundation, vertical integration, as a public policy matter, is typically regarded as having dubious if not outright antisocial properties. Technological interdependencies or, possibly, observational economies, constitute the principal exceptions.

The technological interdependency argument is both the most familiar and the most straight-forward: successive processes which, naturally, follow immediately in time and place dictate certain efficient manufacturing configurations; these, in turn, are believed to have common ownership implications. Such technical complementarity is probably more important in flow process operations (chemicals, metals, etc.) than in separable component

manufacture. The standard example is the integration of iron and steel-making, where thermal economies are said to be available through integration. It is commonly held that where "integration does not have this physical or technical aspect—as it does not, for example, in integrating the production of assorted components with the assembly of those components—the case for cost savings from integration is generally much less clear" (Bain, 1968, p. 381).

There is, nevertheless, a distinct unease over the argument. This is attributable, probably, to a suspicion that the firm is more than a simple efficiency instrument, in the usual scale economies and least-cost factor proportions senses of the term, but also possesses coordinating potential that sometimes transcends that of the market. It is the burden of the present argument that this suspicion is warranted. In more numerous respects than are commonly appreciated, the substitution of internal organization for market exchange is attractive less on account of technological economies associated with production but because of what may be referred to broadly as "transactional failures" in the operation of markets for intermediate goods. This substitution of internal organization for market exchange will be referred to as "internalization."

The two principal prior contributions on which the argument relies are Coase's seminal discussion on "The Nature of The Firm" (1937) and Arrow's more recent review of market versus nonmarket allo-

* Research on this paper has been supported by a grant from the Brookings Institution. It is part of the larger study referred to in footnote 1. Helpful comments from Noel Edelson, Stefano Fenoaltea, Julius Margolis, and Almarin Phillips are gratefully acknowledged.

cation (1969). As will be evident, I agree with Malmgren (1961) that the analysis of transaction costs is uninteresting under fully stationary conditions and that only when the need to make unprogrammed adaptations is introduced does the market versus internal organization issue become engaging.

But while Malmgren finds that the advantage of the firm inheres in its capacity to control information and achieve plan consistency among interdependent activities, which may be regarded as an information processing advantage, I mainly emphasize the differential incentive and control properties of firms in relation to markets. This is not to suggest that information processing considerations are unimportant, but rather that these incompletely characterize the distinctive properties of firms that favor internal organization as a market substitute.

I. *Internal Organization: Affirmative Aspects*

A complete treatment of vertical integration requires that the limits as well as the powers of internal organization be assessed. As the frictions associated with administrative coordination become progressively more severe, recourse to market exchange becomes more attractive, *ceteris paribus*. It is beyond the scope of this paper, however, to examine the organizational failure aspect of the vertical integration question.¹ Rather it is simply asserted that, mainly on account of bounded rationality and greater confidence in the objectivity of market exchange in comparison with bureaucratic processes, market intermediation is generally to be preferred over internal supply in circumstances in

which markets may be said to "work well."²

The properties of the firm that commend internal organization as a market substitute would appear to fall into three categories: incentives, controls, and what may be referred to broadly as "inherent structural advantages." In an incentive sense, internal organization attenuates the aggressive advocacy that epitomizes arms length bargaining. Interests, if not perfectly harmonized, are at least free of representations of a narrowly opportunistic sort; in any viable group, of which the firm is one, the range of admissible intraorganizational behavior is bounded by considerations of alienation. In circumstances, therefore, where protracted bargaining between independent parties to a transaction can reasonably be anticipated, internalization becomes attractive.³

Perhaps the most distinctive advantage of the firm, however, is the wider variety and greater sensitivity of control instruments that are available for enforcing intrafirm in comparison with interfirm activities (Williamson, 1970). Not only does the firm have the constitutional authority and low-cost access to the requisite data which permit it to perform more precise own-performance evaluations (of both a contemporaneous and *ex post* variety)

² An intermediate market will be said to work well if, both presently and prospectively, prices are non-monopolistic and reflect an acceptable risk premium, and if market exchange experiences low transaction costs and permits the realization of essential economies. To the extent that the stipulated conditions do not hold, internal supply becomes relatively more attractive, *ceteris paribus*.

³ Common ownership by itself, of course, does not guarantee goal consistency. A holding company form of organization in which purchaser and supplier are independent divisions, each maximizing individual profits, is no solution. Moreover, merely to stipulate joint profit maximization is not by itself apt to be sufficient. The goal needs to be operationalized, which involves both rulemaking (with respect, for example, to transfer pricing) and the design of efficacious internal incentives. For a discussion, see Williamson (1970).

¹ I discuss the organizational failure dimension of this issue in *Aspects of Monopoly Theory and Policy* (forthcoming). Policy implications of the argument are also examined there.

than can a buyer, but its reward and penalty instruments (which include selective use of employment, promotion, remuneration, and internal resource allocation processes) are more refined.

Especially relevant in this connection is that, when conflicts develop, the firm possesses a comparatively efficient conflict resolution machinery. To illustrate, fiat is frequently a more efficient way to settle minor conflicts (say differences of interpretation) than is haggling or litigation. *Interorganizational* conflict can be settled by fiat only rarely, if at all. For one thing, it would require the parties to agree on an impartial arbitrator, which agreement itself may be costly to secure. It would also require that rules of evidence and procedure be established. If, moreover, the occasion for such interorganizational settlements were to be common, the form of organization converges in effect to vertical integration, with the arbiter becoming a manager in fact if not in name. By contrast, *intraorganizational* settlements by fiat are common (Whinston, 1964, pp. 410-14).

The firm may also resort to internalization on account of economies of information exchange. Some of these may be due to structural differences between firms and markets. Others, however, reduce ultimately to incentive and control differences between internal and market organization. It is widely accepted, for example, that communication with respect to complex matters is facilitated by a common training and experience and if a compact code has developed in the process. Repeated interpersonal interactions may permit even further economies of communication; subtle nuances may come through in familiar circumstances which in an unfamiliar relationship could be achieved only with great effort. Still, the drawing of an organizational boundary need not, by itself, prevent intensely fa-

miliar relations from developing between organizations. Put differently, but for the goal and control differences described above, the informational advantages of internal over market organization are not, in this respect, apparent. Claims of informational economies thus should distinguish between economies that are attributable to information flows *per se* (structure) and those which obtain on account of differential veracity effects (see Part D, Section II).

II. Market Failure Considerations

What are referred to here as market failures are failures only in the limited sense that they involve transaction costs that can be attenuated by substituting internal organization for market exchange. The argument proceeds in five stages. The first three are concerned with characterizing a successively more complex bargaining environment in which small numbers relations obtain. The last two deal with the special structural advantages which, either naturally or because of prevailing institutional rules, the firm enjoys in relation to the market.

A. Static Markets

Consider an industry that produces a multicomponent product, assume that some of these components are specialized (industry specific), and assume further that among these there are components for which the economies of scale in production are large in relation to the market. The market, then, will support only a few efficient sized producers for certain components.

A monopolistic excess of price over cost under market procurement is commonly anticipated in these circumstances—although, as Demsetz (1968) has noted, this need not obtain if there are large numbers of suppliers willing and able to bid at the initial contract award stage. As-

sume, however, that large numbers bidding is not feasible. The postulated conditions then afford an "apparent" incentive for assemblers to integrate backward or suppliers to integrate forward. Two different cases can be distinguished: bilateral monopoly (oligopoly) and competitive assembly with monopolistic supply. The former is considered here; the latter is treated in Part C.

Bilateral monopoly requires that both price and quantity be negotiated. Both parties stand to benefit, naturally, by operating on rather than off the contract curve—which here corresponds to the joint profit maximizing quantity (Fellner, 1947). But this merely establishes the amount to be exchanged. The terms at which this quantity will be traded still need to be determined. Any price consistent with nonnegative profits to both parties is feasible. Bargaining can be expected to ensue. Haggling will presumably continue until the marginal private net benefits are perceived by one of the parties to be zero. Although this haggling is jointly (and socially) unproductive, it constitutes a source of private pecuniary gain. Being, nevertheless, a joint profit drain, an incentive to avoid these costs, if somehow this could be arranged, is set up.

One possible adaptation is to internalize the transaction through vertical integration; but a once-for-all contract might also be negotiated. In a perfectly static environment (one that is free of disturbances of all kinds), these may be regarded with indifference: the former involves settlement on component supply price while merger requires agreement on asset valuation. Bargaining skills will presumably be equally important in each instance (indeed, a component price can be interpreted in asset valuation terms and conversely). Thus, although vertical integration may occur under these conditions, there is nothing in the nature of the prob-

lem that requires such an outcome.

A similar argument in these circumstances also applies to adaptation against externalities: joint profit considerations dictate that the affected parties reach an accommodation, but integration holds no advantage over once-for-all contracts in a perfectly static environment.

Transforming the relationship from one of bilateral monopoly to one of bilateral oligopoly broadens the range of bargaining alternatives, but the case for negotiating a merger agreement in relation to a once-for-all contract is not differentially affected on this account. The static characterization of the problem, apparently, will have to be relaxed if a different result is to be reached.

B. Contractual Incompleteness

Let the above conditions be enriched to include the stipulation that the product in question is technically complex and that periodic redesign and/or volume changes are made in response to changing environmental conditions. Also relax the assumption that large numbers bidding at the initial contract award stage is infeasible. Three alternative supply arrangements can be considered: a once-for-all contract, a series of short-term contracts, and vertical integration.

The dilemma posed by once-for-all contracts is this: lest independent parties interpret contractual ambiguities to their own advantage, which differences can be resolved only by haggling or, ultimately, litigation, contingent supply relations ought exhaustively to be stipulated. But exhaustive stipulation, assuming that it is feasible, is itself costly. Thus although, if production functions were known, appropriate responses to final demand or factor price changes might be deduced, the very costliness of specifying the functions and securing agreement discourages the effort. The problem is made even more severe

where a changing technology poses product redesign issues. Here it is doubtful that, despite great effort and expense, contractual efforts reasonably to comprehend the range of possible outcomes will be successful. An adaptive, sequential decision process is thus indicated. If, however, contractual revisions or amendments are regarded as an occasion to bargain opportunistically, which predictably they will be, the purchaser will defer and accumulate adaptations, if by packaging them in complex combinations their true value can better be disguised; some adaptations may be foregone altogether. The optimal sequential decision-making process can in these respects be distorted.

Short-term contracts, which would facilitate adaptive, sequential decision-making, might therefore be preferred. These pose problems, however, if either (1) efficient supply requires investment in special-purpose, long-life equipment, or (2) the winner of the original contract acquires a cost advantage, say by reason of "first mover" advantages (such as unique location or learning, including the acquisition of undisclosed or proprietary technical and managerial procedures and task-specific labor skills).

The problem with condition (1) is that optimal investment considerations favor the award of a long-term contract so as to permit the supplier confidently to amortize his investment. But, as indicated, long term contracts pose adaptive, sequential decision-making problems. Thus optimal investment and optimal sequential adaptation processes are in conflict in this instance.

It might be argued that condition (2) poses no problems since initial bidders will fully reflect in their original bids all relevant factors. Thus, although anticipated downstream cost advantages (where downstream is used both here and subsequently in the sense of time rather than

place) will give rise to small numbers competition for downstream supply, competition at the initial award stage is sufficient to assure that only competitive returns will be realized over the entire supply interval. One might expect, therefore, that the low bidder would come in at a price below cost in the first period, set price at the level of alternative supply price in later periods, and earn normal returns over-all. Appropriate changes can be introduced easily at the recontracting interval.

A number of potential problems are posed, however. For one thing, unless the total supply requirements are stipulated, "buying in" strategies are risky. Also, and related, the alternative supply price is not independent of the terms that the buyer may subsequently offer to rivals. Moreover, alternative supply price is merely an upper bound; an aggressive buyer may attempt to obtain a price at the level of current costs on each round. Haggling could be expected to ensue. Short-term contracts thus experience what may be serious limitations in circumstances where nontrivial first-mover advantages obtain.

In consideration, therefore, of the problems that both long and short-term contracts are subject to, vertical integration may well be indicated. The conflict between efficient investment and efficient sequential decision-making is thereby avoided. Sequential adaptations become an occasion for cooperative adjustment rather than opportunistic bargaining; risks may be attenuated; differences between successive stages can be resolved more easily by the internal control machinery.

It is relevant to note that the technological interdependency condition involving flow process economies between otherwise separable stages of production is really a special case of the contractual incompleteness argument. The contractual dilemma

is this: On the one hand, it may be prohibitively costly, if not infeasible, to specify contractually the full range of contingencies and stipulate appropriate responses between stages. On the other hand, if the contract is seriously incomplete in these respects but, once the original negotiations are settled, the contracting parties are locked into a bilateral exchange, the divergent interests between the parties will predictably lead to individually opportunistic behavior and joint losses. The advantages of integration thus are not that technological (flow process) economies are unavailable to nonintegrated firms, but that integration harmonizes interests (or reconciles differences, often by fiat) and permits an efficient (adaptive, sequential) decision process to be utilized. More generally, arguments favorable to integration that turn on "supply reliability" considerations commonly reduce to the contractual incompleteness issue.⁴

C. Strategic Misrepresentation Risk

Contractual incompleteness problems develop where there is *ex ante* but not necessarily *ex post* uncertainty. Strategic misrepresentation risks are serious where there is uncertainty in both respects. Not only is the future uncertain but it may not be possible, except at great cost, for an outside agency to establish accurately what has transpired after the fact. The advantages of internalization reside in the

facts that the firm's *ex post* access to the relevant data is superior, it attenuates the incentives to exploit uncertainty opportunistically, and the control machinery that the firm is able to activate is more selective.

1. AFFIRMATIVE OCCASIONS for INTEGRATION. Three affirmative occasions to integrate on account of strategic misrepresentation risk and two potentially anticompetitive consequences of integration can be identified.

(a) MORAL HAZARD. The problem here arises because of the conjoining of inharmonious incentives with uncertainty—or, as Arrow puts it (1969, p. 55), it is due to the "confounding of risks and decisions." To illustrate, consider the problem of contracting for an item the final cost and/or performance of which is subject to uncertainty. One possibility is for the supplier to bear the uncertainty. But, he will undertake a fixed price contract to deliver a specified result the costs of which are highly uncertain only after attaching a risk premium to the price. Assume that the buyer regards this premium as excessive and is prepared on this account to bear the risk himself. The risk can easily be shifted by offering a cost-plus contract. But this impairs the incentives of the supplier to achieve least-cost performance; the supplier may reallocate his assets in such a way as to favor other work to the disadvantage of the cost-plus contract.

Thus, although, if commitments were self-enforcing, it might often be institutionally most efficient to divide the functions of risk bearing and contract execution (that is, cost-plus contracts would have ideal properties), specialization is discouraged by interest disparities. At a minimum, the buyer may insist on monitoring the supplier's work. In contrast therefore to a fixed-price contract, where it is sufficient to evaluate end-product performance, cost-plus contracts, because

⁴It is sometimes suggested that breach of contract risk affords an additional reason for integration: the small supplier of a critical component whose assets are insufficient to cover a total damage claim leaves the purchaser vulnerable. But this is an argument against small suppliers, not contracting quite generally; the large, diversified supplier might well have superior risk pooling capability to that of the integrated firm. The risks of contractual incompleteness, however, remain and may discourage purchasing from large, diversified organizations. For a discussion of "ideal" contracts in this connection, see Arrow (1965, pp. 52–53).

they expose the buyer to risks of inefficient (high cost) contract execution, require that *both* inputs and outputs be evaluated.

Internalization does not eliminate the need for input evaluation. Rather, the advantage of internalization, for input monitoring purposes, resides in the differential ease with which controls are exercised. An external agency, by design, lacks recourse to the internal control machinery: proposed remedies require the consent of the contractor and then are highly circumscribed; unrestricted access by the buyer to the contractor's internal control machinery (including selective use of employment, promotion, remuneration, and internal resource allocation processes) is apt to be denied. In consideration of the costs and limitations of input monitoring by outsiders, the buyer may choose instead to bear the risk and perform the work himself. The buyer thus internalizes, through backward vertical integration, a transaction which, but for uncertainty, would move through the market. A cost-type contract for *internal procurement* is arranged.

(b) EXTERNALITIES/IMPUTATION. The externality issue can be examined in two parts. First, has a secure, unambiguous, and "appropriate" assignment of property rights been made? Second, are the accounting costs of imputing costs and benefits substantial? If answers to these questions are affirmative and negative respectively, appropriability problems will not become an occasion for vertical integration. Where these conditions are not satisfied, however, integration may be indicated.

The assignment aspect of this matter is considered in Part E below. Here it is assumed that an efficacious assignment of property rights has been made and that only the expense of imputing costs and benefits is at issue. But indeed this is apt often to be the more serious problem.

High imputation expenses which discourage accurate metering introduce ambiguity into transactions. Did party A affect party B and if so in what degree? In the absence of objective, low cost standards, opposed interests can be expected to evaluate these effects differently. Internalization, which permits protracted (and costly) disputes over these issues to be avoided, may on this account be indicated.

(c) VARIABLE PROPORTIONS: DISTORTIONS. Consider the case where the assembly stage will support large numbers; fewness appears only in component supply. Whether monopolistic supply prices provide an occasion for vertical integration in these circumstances depends both on production technology and policing expense. Variable proportions at the assembly stage afford opportunities for nonintegrated assemblers to adapt against monopolistically priced components by substituting competitively priced factors (McKenzie, 1951). Although conceivably the monopolistic component supplier could stipulate, as a condition of sale, that fixed proportions in assembly should prevail, the effectiveness of such stipulations is to be questioned—since, ordinarily, the implied enforcement costs will be great. Where substitution occurs, inefficient factor proportions, with consequent welfare losses, will result. The private (and social) incentives to integrate so as to reduce total costs by restoring efficient factor combinations are evident.

2. ANTICOMPETITIVE CONSEQUENCES. Anticompetitive effects of two types are commonly attributed to integration: price discrimination and barriers to entry (cf. Stigler, 1968, p. 303).

(a) PRICE DISCRIMINATION. The problem here is first to discover differential demand elasticities, and secondly to arrange for sale in such a way as to preclude reselling. Users with highly elastic demands which purchase the item at a low price

must not be able to service inelastic demand customers by acting as a middleman; all sales must be final. Although vertical integration may facilitate the discovery of differential elasticities, it is mainly with respect to the non-resale condition that it is regarded as especially efficacious.

Integration, nevertheless, is a relatively extreme response. Moreover, price discrimination is clearly practiced in some commodities without recourse to vertical integration (witness electricity and telephone service). What are the distinguishing factors? Legality considerations aside, presumably it is the cost of enforcing (policing) terms of the contract that are at issue. Some commodities apparently have self-enforcing properties—which may obtain on account of high storage and re-packing costs or because reselling can not be arranged inconspicuously. The absence of self-enforcing (policing) properties is what makes vertical integration attractive as a means of accomplishing discrimination.

(b) ENTRY BARRIER EFFECTS. That the vertical integration of production might be used effectively to bar entry is widely disputed. Bork (1969, p. 148) argues that "In general, if greater than competitive profits are to be made in an industry, entry should occur whether the entrant has to come in at both levels at once or not. I know of no theory of imperfections in the capital market which would lead suppliers of capital to avoid areas of higher return to seek areas of lower return." But the issue is not one of profit avoidance but rather involves cost incidence. If borrowers are confronted by increasingly adverse rates as they increase their finance requirements, which Hirshleifer suggests is a distinct possibility (1970, pp. 200-1), cost may not be independent of vertical structure.

Assuming that vertical integration has the effect of increasing capital requirements, the critical issues are to what ex-

tent and for what reasons the supply curve of finance behaves in the way postulated. The following conjecture is offered as a partial explanation: unable to monitor the performance of *large, complex* organizations in any but the crudest way or to effect management displacement easily except on evidence of seriously discreditable error, investors demand larger returns as finance requirements become progressively greater, *ceteris paribus*. Thus the costs of policing against the contingency that managers will operate a rival enterprise opportunistically are, on this argument, at least partly responsible for the reputed behavior of the supply curve of capital. In consideration of this state of affairs, established firms may use vertical integration strategically to increase finance requirements and thereby to discourage entry if potential entrants feel compelled, as a condition of successful entry, to adopt the prevailing structure—as they may if the industry is highly concentrated.

D. Information Processing Effects

As indicated in Section I, one of the advantages of the firm is that it realizes economies of information exchange. These may manifest themselves as information impactedness, observational economies, or what Malmgren (1961) refers to as the "convergence of expectations."

1. INFORMATION IMPACTEDNESS. Richardson illustrates the problems of information impactedness by reference to an entrepreneur who was willing to offer long-term contracts (at normal rates of return, presumably) but which contracts others were unprepared to accept because they were not convinced that he had "the ability, as well as the will, to fulfill them. He may have information sufficient to convince himself that this is the case, but others may not" (Richardson, 1960, p. 83). He goes on to observe that the perceived

risks of the two parties may be such as to make it difficult to negotiate a contract that offers commensurate returns to each; objective risks are augmented by contractual risks in these circumstances. Integration undertaken for this reason is akin to self-insurance by individuals who know themselves to be good risks but are priced out of the insurance market because of their inability, at low cost, to "reveal" this condition to insurers.

2. OBSERVATIONAL ECONOMIES. As Radner indicates, "the acquisition of information often involves a 'set-up cost'; i.e., the resources needed to obtain the information may be independent of the scale of the production process in which the information is used" (Radner, 1970, p. 457). Although Radner apparently had horizontal firm size implications in mind, the argument also has relevance for vertical integration. If a single set of observations can be made that is of relevance to a related series of production stages, vertical integration may be efficient.

Still, the question might be raised, why common ownership? Why not an independent observational agency that sells information to all comers? Or, if the needed information is highly specialized, why not a joint venture? Alternatively, what inhibits efficient information exchange between successive stages of production according to contract? In relation, certainly, to the range of intermediate options potentially available, common ownership appears to be an extreme response. What are the factors which favor this outcome?

One of the problems with contracts is that of specifying terms. But even if terms could be reached, there is still a problem of policing the agreement. To illustrate, suppose that the common information collection responsibilities are assigned by contract to one of the parties. The purchasing party then runs a *veracity risk*: information may be filtered and possibly

distorted to the advantage of the firm that has assumed the information collection responsibility. If checks are costly and proof of contractual violation difficult, contractual sharing arrangements manifestly experience short-run limitations. If, in addition, small numbers prevail so that options are restricted, contractual sharing is subject to long-run risks as well. On this argument, observational economies are mainly to be attributed to strategic misrepresentation risks rather than to indivisibilities.

3. CONVERGENCE OF EXPECTATIONS. The issue to which the convergence of expectations argument is addressed is that, if there is a high degree of interdependence among successive stages of production and if occasions for adaptation are unpredictable yet common, coordinated responses may be difficult to secure if the separate stages are operated independently. March and Simon (1958, p. 159) characterize the problem in the following terms:

Interdependence by itself does not cause difficulty if the pattern of interdependence is stable and fixed. For, in this case, each subprogram can be designed to take account of all the subprograms with which it interacts. Difficulties arise only if program execution rests on contingencies that cannot be predicted perfectly in advance. In this case, coordinating activity is required to secure agreement about the estimates that will be used as the basis for action, or to provide information to each subprogram unit about the activities of the others.

This reduces, in some respects, to a contractual incompleteness argument. Were it feasible exhaustively to stipulate the appropriate conditional responses, coordination could proceed by contract. This is ambitious, however; in the face of a highly variable and uncertain environment, the attempt to program responses is apt to be inefficient. To the extent that an unprogrammed (adaptive, sequential)

decision process is employed instead, and in consideration of the severe incentive and control limitations that long-term contracts experience in these circumstances (See Part B above), vertical integration may be indicated.

But what of the possibility of short-term contracts? It is here that the convergence of expectations argument is of special importance. Thus assume that short-term contracts are not defective on account either of investment disincentives or first-mover advantages. It is Malmgren's (1961) contention that such contracts may nevertheless be vitiated by the absence of structural constraints. The costs of negotiations and the time required to bring the system into adjustment by exclusive reliance on market (price) signals are apt to be great in relation to that which would obtain if successive states were integrated and administrative processes employed as well or instead.

E. Institutional Adaptations

Institutional adaptations of two types are distinguished: simple economic and extra-economic.

1. **SIMPLE ECONOMIC.** As has been noted by others, vertical integration may be a device by which sales taxes on intermediate products are avoided, or a means by which to circumvent quota schemes and price controls (Coase, 1937, pp. 338-39; Stigler, 1968, pp. 136-37). But vertical intergration may also be undertaken because of the defective specification of property rights.

Although the appropriate assignment of property rights is a complex question, it reduces (equity considerations aside) to a simple criterion: What assignment yields maximum total product (Coase, 1960, p. 34)? This depends jointly on imputation and negotiation expenses and on the incentives of the compensated party. So as to focus on the negotiation expense aspect,

assume that imputation expenses are negligible and set the incentive question aside for the moment.⁵ An "appropriate" assignment of property rights will here be defined as one which automatically yields compensation in the amount of the external benefit or cost involved, while an "inappropriate" assignment is one that requires bargaining to bring the parties into adjustment. Thus if A and B are two parties and A's activity imposes costs on B, the appropriate assignment of property rights is to require A to compensate B. If instead property rights were defined such that A is not required to compensate B, and assuming that the externality holds at the margin, efficient adaptation would occur only if B were to bribe A to bring his activity into adjustment—which entails bargaining. Only if the costs of such bargaining are neglected can the alternative specifications of property rights be said to be equivalent. For similar reasons, if A's activity generates benefits for B, the appropriate specification of property rights will be to require B fully to compensate A. Harmonizing the otherwise divergent interests of the two parties by internalizing the transaction through vertical merger

⁵ As Coase has emphasized (1960, pp. 32-33, 41), compensation can impair the incentives of the compensated party that experiences an external cost to take appropriate protective measures. Parties that are assured of compensation will be content to conduct business as usual. Such a practice easily contributes to greater social cost than would obtain were compensation denied. A sensitivity to what, in a broad sense, might be regarded as contributory negligence is thus required if the system is to be brought fully into adjustment. Clairvoyance with respect to contributory negligence would of course permit the courts to supply those who experience the external cost with requisite incentives to adapt appropriately. Since, however, such clairvoyance (or even unbiasedness) cannot routinely be presumed, internalizing the transaction through vertical integration may be indicated for this reason as well. (Interestingly, a symmetrical problem is not faced where the externality is a benefit. Stipulating that compensation shall be paid induces Meade's (1952) orchard grower not merely to extend his production appropriately, but also to shift from apples to peaches if this is socially advantageous.)

promises to overcome the haggling costs which result when property rights are left either undefined or inappropriately specified.

2. **OTHER.** Risk aversion refers to the degree of concavity in the utility valuation of pecuniary outcomes. Decision-makers who are risk averse will be concerned not merely with the expected value, but also with the dispersion in outcomes associated with alternative proposals: the greater the dispersion, the lower the utility valuation. *Ceteris paribus*, decision-makers who are the less risk averse will presumably assume the risk bearing function. Even, however, if attitudes toward risk were identical—in the sense that every individual (for any given set of initial endowments) would evaluate a proposal similarly—differing initial asset positions among the members of a population could warrant a specialization of the risk-bearing function, with possible firm and market structure effects (Knight, 1965).

Arrow calls attention to norms of social behavior, including ethical and moral codes. He observes in this connection that "It is useful for individuals to have some trust in each other's word. In the absence of trust, it would become very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be foregone" (1969, p. 62). One would expect, accordingly, that vertical integration would be more complete in a low-trust than a high-trust culture, *ceteris paribus*.

III. Conclusions

That product markets have remarkable coordinating properties is, among economists at least, a secure proposition. That product markets are subject to failure in various respects and that internal organization may be substituted against the

market in these circumstances is, if somewhat less familiar, scarcely novel. A systematic treatment of market failure as it bears on vertical integration, however, has not emerged.

Partly this is attributable to inattention to internal organization: the remarkable properties of firms that distinguish internal from market coordination have been neglected. But the fragmented nature of the market failure literature as it bears on vertical integration has also contributed to this condition; the extensive variety of circumstances in which internalization is attractive tends not to be fully appreciated.

The present effort attempts both to address the internal organization issue and to organize the market failure literature as it relates to vertical integration in a systematic way. The argument, however, by no means exhausts the issues that vertical integration raises. For one thing, the discussion of market failures may be incomplete in certain respects. For another, a parallel treatment of the sources and consequences of the failures of internal organization as they relate to vertical integration is needed. Third, the argument applies strictly to the vertical integration of production; although much of it may have equal relevance to backward vertical integration into raw materials and forward integration into distribution, it may have to be delimited in significant respects. Fourth, game theoretic considerations, which may permit the indicated indeterminacy of small numbers bargaining situations to be bounded, have been neglected. Finally, nothing in the present analysis establishes that observed degrees of vertical integration are not, from a social welfare standpoint, excessive. It should nevertheless be apparent that a broader *a priori* case for the vertical integration of production exists than is commonly acknowledged.

REFERENCES

- Kenneth J. Arrow, *Aspects of the Theory of Risk-Bearing*, Helsinki, 1965.
- , "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation," in *The Analysis and Evaluation of Public Expenditures: The PPB System*, Vol. 1, Joint Economic Committee, Washington, D.C., 1969, pp. 47-64.
- Joe S. Bain, *Industrial Organization*, New York, 1968.
- Robert H. Bork, "Vertical Integration and Competitive Processes," in J. Fred Weston and Sam Peltzman, eds., *Public Policy Toward Mergers*, Pacific Palisades, Calif., 1969, pp. 139-49.
- Ronald H. Coase, "The Nature of the Firm," *Economica*, Nov. 1937, 4, 386-405; reprinted in George J. Stigler and Kenneth E. Boulding, eds., *Readings in Price Theory*, Homewood, Ill., 1952, pp. 331-51.
- , "The Problem of Social Cost," *Journal of Law and Economics*, 1960, 3, 1-44.
- Harold Demsetz, "Why Regulate Utilities?" *Journal of Law and Economics*, April 1968, 11, 55-66.
- William Fellner, "Prices and Wages under Bilateral Oligopoly," *Quarterly Journal of Economics*, August 1947, 61, 503-32.
- J. Hirshleifer, *Investment, Interest and Capital*, Englewood Cliffs, N.J., 1970.
- Frank H. Knight, *Risk, Uncertainty and Profit*, New York, 1965.
- Lionel McKenzie, "Ideal Output and the Interdependence of Firms," *Economic Journal*, December 1951, 61, 785-803.
- H. B. Malmgren, "Information, Expectations and the Theory of the Firm," *Quarterly Journal of Economics*, August 1961, 75, 399-421.
- James E. Meade, "External Economies and Diseconomies in a Competitive Situation," *Economic Journal*, March 1952, 62, 54-67.
- Roy Radner, "Problems in the Theory of Markets under Uncertainty," *American Economic Review*, May 1970, 60, 454-60.
- G. B. Richardson, *Information and Investment*, London, 1960.
- George J. Stigler, *The Organization of Industry*, Homewood, Ill., 1968.
- Andrew Whinston, "Price Guides in Decentralized Organizations," in W. W. Cooper, H. J. Leavitt, and M. W. Shelly, II, eds., *New Perspectives in Organization Research*, New York, 1964, pp. 405-48.
- Oliver E. Williamson, *Corporate Control and Business Behavior*, Englewood Cliffs, N.J., 1970.

DISCUSSION

ROLAND N. MCKEAN: Williamson is undoubtedly correct, in my view, in urging that an important reason for vertical integration is the attempt to reduce transaction costs. Hagglng, explicit contracting, and enforcement are costly activities. I am much impressed by the extent to which social intercourse and markets themselves rely on trust, tacitly understood contracts, and, for enforcement, latent social or retaliatory pressures. Transaction costs can explain much behavior that has been unexplained and can correct many explanations that have been at least partly wrong.

My only criticisms are that Williamson might have helped the reader by (1) showing more fully how the various points fit together and (2) looking more consistently at property rights and thus probing further into the reasons for the behavior he describes.

(1) The following way of presenting or summing up the ideas might help one see the way Williamson's points fit together. Without vertical integration there are deals that would be mutually advantageous at low transaction costs but are not carried out because of high transaction costs, and other deals that are carried out but at unnecessarily high transaction costs. These potential or actual deals include information exchange, product purchase, redistribution of risk bearing, elimination of inefficient input combinations by the processor because of monopoly prices charged by component manufacturers, and achievement of technological economies by arranging lower-cost transfer of components to the processor (e.g., molten iron to the steel mill). All of these deals would amount to internalizing externalities or taking interdependencies into account, so I see no reason for a separate category of that sort. They are all externalities because in each instance there is a cost being imposed on, or a potential benefit being denied to, one firm or the other without contractual agreement. Moreover, these interdependencies, i.e., ones involving high transaction costs, are especially pervasive in a dynamic world involving product redesign, durable investments, and numerous contingencies. (If there were zero transaction costs, it might be noted, all

these potentially advantageous arrangements would be made—and, needless to say, at minimum transaction cost.)

Next one could explain the various ways in which vertical integration, up to a point, can reduce these transaction costs, making additional exchanges economical and making the old ones possible at lower cost.

(2) One could clarify this latter explanation by probing into the claim- or right-structures of the managers and other personnel involved. Why is haggling less prevalent (it is by no means absent!) within an integrated firm than between independent components producers and a processor? Basically, it is because the manager of an independent firm has a claim on part of *that* firm's profits, and lower-level personnel in turn have claims on (i.e., they are rewarded for contributing to) *that* firm's profits. If these personnel work instead for the integrated firm, they are rewarded (more than before, at least) for increasing the integrated firm's profits. The instruments of control mentioned by Williamson—hiring, firing, promotion, salary changes, transfer of personnel—amount to a reassignment of certain rights to the integrated firm. These instruments, or rights over the new employees, work to alter those employees' claims as noted and to enforce them.

Because of this altered claim structure, it is now less rewarding to management of the components division to haggle with the management of the processing division (and vice versa). It is less rewarding than before to exchange dubious information or to engage in strategic bargaining. It becomes more rewarding than before for personnel in all divisions to accept monitoring by overall management. Employees find that flats by the new management are less costly and/or more rewarding to them than when their rights were linked to the profits of different firms. (All this is obvious, but the point is that it is not merely the "conflict-resolving machinery" or the "mechanisms" that have changed but also the effective claims held by the individuals involved.) Now that they are rewarded for veracity to the new management and contributions to

overall profits, the divisions can trust each other and work harmoniously to a greater extent. Informal, tacitly understood contracts become enforceable and are efficient more often than before.

I do not mean that the changes in claim structures are the *only* variables at work. Contracting, information flows, or monitoring *may* become physically easier with integration. Such impacts are not obvious to me, however. The physical location of personnel and materials, and the physical flows, *could* have been adjusted before. As Williamson points out, the so-called technological economies were physically attainable all along; the obstacle must have been transaction costs. I would add that any economies in contracting, information acquisition, or monitoring were also physically attainable all along; the obstacle must have been something about incentives. In any event, I merely suggest that changes in rights with vertical integration constitute a major factor in the reduction of transaction costs. Even Williamson's suggestion that "common experience" (after a merger) may make it easier to communicate seems to hinge on cost-reward patterns, for typically the components people would still work in a separate location and not really have much additional "common experience." And when he writes that "the range of admissible intraorganizational behavior is bounded by considerations of alienation," one of the main *extra* reasons for avoiding alienation is the possible impact on promotion, salary, and so on. Even low-cost access to data is a matter of incentive: the files existed before integration and will probably be in the same location after integration, but the manager of the components division now finds it rewarding to give top management access to his files.

Looking at these underlying claims is instructive also in that it suggests caution in appraising other organizational changes. If various factors, perhaps even bigness, attenuate the claims of personnel to a "share of profits," intrafirm transaction costs will go up. Also, where goals are diverse and claims are not linked to increased profits—as in government agencies—it does not follow that integration or conglomeration will markedly reduced trans-

action costs. It is also interesting to speculate about possible trends. If technological or other developments make information more valuable, externalities between firms more important, or internal monitoring less expensive, one would presumably expect the readjustment of incentives via integration to become more attractive and more frequent.

J. FRED WESTON: The first part of Professor Lintner's paper is a summary of four major-merger movements. His judgments of the relative importance of three motivating influences are summarized in Table 1.

TABLE 1—ANALYSIS OF FOUR MAJOR-MERGER MOVEMENTS

Relative Importance of Motivating Influences

Merger Movements	Market Power (h)	Operating Economies (s)	Leakages to Promoters (b)
Turn of the Century	large	small	large
1920's	moderate	moderate	moderate
1941-1947	small	moderate	small
Post 1950	small	small	small

The *effects* of the merger movements may have differed from the motives ascribed by Professor Lintner. The nearly 100% market shares of dominant firms at the turn of the century declined substantially and almost without exception. Thus the market position objectives motivating the mergers were not maintained in the ensuing decades. The oligopolistic market structures developed during the 1920's resulted in market share and profit trends consistent with operating economies rather than market power. Consequently the expected gains from mergers capitalized into increased market prices of securities were not generally realized as documented by the studies Lintner cites. Competition eroded market positions and imitation reduced differential operating efficiencies.

The second part of Professor Lintner's paper analyzed the post-1950 conglomerate merger movement. He evaluates potential sources of gains from the criterion of private advantage measured by increases in immedi-

TABLE 2—POTENTIAL SOURCES OF GAINS IN CONGLOMERATE MERGERS UNDER THE ASSUMPTIONS THAT MARKET POWER AND OPERATING ECONOMIES ARE ZERO.

Private Gain	Social Gain	Potential Sources of Gain
yes	no	1. Stimulus to "imaginative accounting"
		2. Tax benefits
yes	no	a. loss carry overs
yes	no	b. tax deferrals
		3. Leverage
yes	no	a. reduction of under-utilized borrowing capacity
yes	yes	b. economies of scale in financing
yes	yes	c. reduction of joint probability of bankruptcy to the firm
		4. All equity case
yes	no	a. Unequal price/earnings ratios
yes	yes	b. Revise assessments of future values
		(1). efficient generation of information
		(2). reduced calculation costs
yes	yes	c. Reduce the market price of risk by increasing the number of holders of a security

ate market values of securities. I shall summarize Lintner's findings, adding the criterion of social gains, and relate his theorems to empirical data.

Professor Lintner deemphasizes the generalizations summarized from the literature in Table 2, lines 1 to 3a. In addition, I would observe that the Levy-Sarnat, Lewellen theorem (line 3c) requires that both the probabilities and costs of bankruptcy be of material magnitudes.

The last two items under the all equity case, represent original contributions of Professor Lintner. They represent extensions of his earlier theorems on capital asset pricing. Here the issue is the applicability of the models, given their assumptions.

A number of empirical studies on conglomerates are consistent with Professor Lintner's theorems. Westerfield measured one aspect of the efficiency of diversification by the ratio of systematic to total risk (i.e., the elimination of unsystematic risk by diversification). He

found that conglomerates ranked low in comparison with mutual funds, predictable by the theory of conglomerates since investors can diversify directly and efficiently.

The Smith and Schreiner study also found that the extent of diversification by conglomerate firms was small relative to mutual funds. However, they also measured the efficiency of diversification, employing a simulation technique using the investment returns of the Standard and Poors' industry categories. Their performance test was Sharpe's initial measure, risk premium divided by the standard deviation of returns. The top four were mutual funds, but thereafter an intermixture in rankings between conglomerates and mutual funds was observed.

In an extension of this study Smith, Weston, and Shrieves utilized *actual* performance over the period 1960-1969. Some conglomerates outranked all mutual funds indicating that the potentials described by Lintner can in fact be realized. When the Treynor measure (later also adopted by Sharpe) is used (dividing the risk premium by "volatility"), the results are similar. This finding supports Professor Lintner's prediction of the favorable effects of conglomerate mergers on the market price of risk.

An alternative theory of conglomerates was presented by Dennis Mueller. In Mueller's theory, the conglomerate movement resulted from the separation of ownership and management control in larger, mature corporations. If the rewards to managers are a function of the size of the firm, managers are motivated to maximize growth in the size of firms, utilizing a lower investment hurdle rate.

In a study critical of earlier evidence, Lewellen and Hunstman present findings that managers' compensation is significantly correlated with the profit rate, not the rate of sales of firms. Thus the premise of the Mueller theory is doubtful.

In a later exchange, Mueller rejects evidence of higher average returns in manager vs. owner controlled firms, arguing that only marginal returns are relevant, which implies that average returns should fall over time. Earnings performance data on a sample of 63 conglomerate firms bears directly on these issues.

Utilizing measures relatively free of the influence of "merger accounting," the rate of earnings plus interest but after taxes to total assets for the conglomerates in 1958 (at the beginning of the movement) was 5.8%, significantly lower than the 9.2% return for a random sample of non-conglomerate firms. By 1968 or 1969, the difference was nonsignificant, indicating that the conglomerates had moved from below average earning performance to average earnings performance. These findings are inconsistent with Mueller's theory, but support Lintner's theorems with respect to conglomerate firms; in addition, the data indicate that his "s" factor may also have been appreciable.

This raises the question of the theoretical basis for a positive "s" factor. Here Williamson's brilliant exposition of the role of incentives and control systems in extending Coase's theory of the firm is applicable. I would go beyond Williamson's position to argue that organization quality deserves explicit recognition in the specification of the neoclassical production function of the firm. Three considerations are relevant: (1) The time variable is important in organization quality. In this connection, I am testing the correlation between firm performance and its management organization continuity. (2) The unequal steepness of organization learning curves has an influence on performance results over time. (3) The effects of differential organizational quality have cumulative impacts. Just as the cumulative effects of differential advertising may result in a capital requirements entry barrier, the cumulative

effects of differential organizational quality may result in an efficiency barrier to entry.

These propositions help explain the acceleration of the conglomerate movement after the mid-1950's when a revolution in managerial technology occurred. Enterprise planning began to develop and advances in computer technology began to be adapted to management of the firm. Financial planning and control systems were extended with further improvements in the use of balanced centralized-decentralized management control systems. Further, World War II and the Korean Conflict had stimulated new technologies resulting in an uneven diffusion and wide variations in advanced technological capabilities among firms. The major conceptual point here is that the role of the general management functions (planning, control, organizing, information systems) and functions centralized at top management levels (research, finance, legal) increased in importance in the management of enterprise. In consequence, the costs of managing large, diversified firms were substantially reduced relative to potential operating economies.

This is the broader theoretical basis explaining the ability of the conglomerate firms as a group to raise their average return on economic resources during the decade of the 1960's. Professor Lintner demonstrates that even in the absence of a positive "h" or "s" factor, gains from conglomerate mergers exist. My research suggests a positive "s" factor as well.

MICRO ASPECTS OF MACRO PERFORMANCE

The Impact of Inflation on the Distribution of Income and Wealth

By EDWARD C. BUDD

The Pennsylvania State University and Office of Business Economics

AND DAVID F. SEIDERS

Board of Governors of the Federal Reserve System

One of the major charges against inflation is that it results in arbitrary changes in the distribution of income and wealth among individuals and groups in the economy. Despite the fact that such distributional effects are an important public policy issue, the literature on the topic is surprisingly thin. Most of the studies have been limited to a consideration of either sectoral effects or effects upon groups defined along worker-capitalist or debtor-creditor lines, with little attention to the more general and socially relevant question of the distribution of the effects of inflation among families or individuals classified by size of income or net worth. This study is an attempt to fill part of this gap.

We have developed a model which permits us to estimate, through the use of a micro data file, the effect of inflation (defined as a rise in the price level of goods and services) on the size distribution of real income and net worth. The first step in its construction was the theoretical analysis and empirical estimation, from time series data from the post-World War II period in the U.S., of the effects of inflation on the real value of various asset, debt, and income *types* (as listed in Tables 1 and 2). In the second step, the price level elasticities or adjustment coefficients obtained in the first step were applied to the individual income, asset, and/or debt

TABLE 1—PRICE ADJUSTMENT COEFFICIENTS
FOR NET WORTH MODEL

Net Worth Component	Adjustment Coefficient
Publicly traded common stock, including mutual funds	1.118
Equity in closely-held corporations	1.207
Equity in nonfarm proprietorships and partnerships	1.144
Equity in farm business	1.122
Investment real estate, owner-occupied homes, equity in profit sharing plans, and "other assets" (royalties, patents, etc.)	1.00
Claims (checking, saving, and brokerage accounts; bonds, mortgages, notes, and other loans; preferred stock)	0.00
Debts (mortgages, loans, and noninstallment debt; debit balance on brokerage accounts)	0.00

Source: Methodological Appendix

types received or held by each consumer unit, as reported in the micro data file, in a simulated inflation based on an assumed rate of price increase. In the third step, consumer units were ranked before and after the simulated inflation by size of income or net worth, divided into quantiles of various size (e.g., percentiles, deciles), and the percentage changes in real income or net worth for the quantile or in its share in them were computed. The effect of the simulated inflation on different parts of the size distribution thus can easily be determined.

The data file used was the Survey of Financial Characteristics of Consumers (SFCC), conducted by the Bureau of the Census for the Federal Reserve Board for income during 1962 and asset holdings and liabilities as of the end of 1962. This survey provides considerably more detail on income types than is available from other sample surveys, such as the Current Population Survey, and in addition contains quite detailed data on individual asset and liability types. The sample's stratification by income level permits more reliable estimates of the upper tail of the

distribution, especially important for asset holdings which are concentrated at the top, although the smaller sample sizes in the lower brackets preclude quantile groupings defined as narrowly as for the top. A comprehensive discussion of the survey and the sample, as well as sampling variability and response and nonresponse error, is given in [3, pp. 45-62] and [4, pp. 113-21].

It should be noted at the outset that the problem analyzed is more limited than the broad, and probably unmanageable, one, "How is the distribution of income and wealth affected by inflation?" Our problem can be phrased as follows: given the historical evidence on the extent to which economic decisions have become adjusted to inflation, how would a change in the rate of inflation affect the size distribution of income and wealth? The role of expectations and resulting lags in adjustment is vital. Indeed, if the rate of inflation has been fully anticipated, reflected in individual behavior and hence in market prices and rates of return, and embodied in the appropriate contracts—such as in current rates of interest and in wage escalator clauses—the realization of the anticipated rate need not result in any redistribution, and we would be left without a problem to analyze. Second, our model is more concerned with the effect of a change in the rate of inflation on the distribution in the first (annual) period in which it occurs, rather than the ultimate effect of a constant rate of price increase. Our model could supply a tentative answer to the latter question, too, although the experience of such a constant rate over many periods might serve to alter the empirically derived price elasticities on which our estimates are based. On the other hand, there is nothing particularly sacred about the changes in rates of inflation assumed in our simulations (2% and 5%) or the level from which these changes are assumed to

TABLE 2—PRICE ADJUSTMENT COEFFICIENTS
FOR INCOME MODELS

Income Component	Adjustment Coefficient
Wage and salary income	1.000
Income of unincorporated enterprises (farm and nonfarm)	1.037
Dividends from publicly-held corporations	0.854
Income (after tax) from closely-held corporations	-0.757
Interest on U. S. government bills	6.790
Interest on other U. S. government marketable securities	0.775
Interest on U. S. savings bonds	0.183
Interest on state and local securities	0.327
Interest on corporate and foreign securities	0.178
Interest on mortgages and other loans	0.294
Interest on time deposits of commercial banks	0.698
Interest on savings accounts, savings and loan associations	0.280
Interest on savings accounts, mutual savings banks	0.370
Interest on credit union shares	0.148
Net rental income	0.230
Royalty income	1.000
Transfer income:	
Pensions and annuities, $p = .05$	0.620
Pensions and annuities, $p = .02$	0.390
Other income	0.000
Income items imputed to consumer units as owners of corporations:	
Retained earnings, publicly-held corporations	-1.916
Profits before tax, publicly-held corporations	1.074
Profits before tax, closely-held corporations	1.115

Source: Methodological Appendix

occur (a constant price level). In view of the empirical estimates, however, it would not be fair to use a much larger rate than those experienced in the past; in any case, the model applies to "creeping" inflation of the U.S. variety rather than "galloping" or hyperinflation.

Third, the model abstracts from changes in the level of real output and employment that may be associated with changes in the rate of inflation; our distributional effects could be viewed, for instance, as a consequence of a change in the rate of inflation at a given level of unemployment, associated, say, with an upward shift of the Phillips curve. There may, of course, be a reduction in unemployment accompanying an increase in aggregate demand, which could be represented by a movement along a given Phillips curve to a higher rate of inflation. The consequent increase in real output would itself have distributional effects, e.g., an increased profits share and possibly some narrowing of wage differentials, as well as the increased share of wage income to those previously unemployed. The effect of any inflationary movement, however classified, on distribution can be viewed as a combination of both price and real output effects. Estimation of the latter is beyond the scope of this paper and would require an extensive research project all its own.

Fourth, the model abstracts from differential movements of price indices relevant to the expenditures of different economic groups in the economy. While the effects of such differential movements could be incorporated if the relevant price indices were available, they are probably of minor importance. A recent study [2] involving construction of price indices relevant to the expenditure patterns of poor and near poor, for example, suggests that differential effects on the expenditure side are relatively small.

Fifth, we are interested primarily in the effect of inflation on the distribution of income as determined by market forces and by public and private transfer policies, before any subsequent redistribution through the personal income tax. Strictly speaking, for a before-tax distribution we should treat the corporate income tax in the same way as the personal income tax—as part of the income of shareholders—although this possibility is precluded by the definition of money income used in the SFCC, which includes in the income of shareholders of publicly-held corporations only dividends received. Two of our alternative concepts, which allocate corporate retained earnings, and retained earnings plus corporate income taxes, on the basis of dividends received, are designed to show approximately how much difference the exclusion of these two components of corporate profits from the money income concept makes to our results.

Finally, the effect of inflation on each income or asset type, in the full detail reported in the SFCC, is assumed to be uniform for each recipient or holder of that type. While it might be possible to develop techniques to achieve further breakdowns of SFCC asset and income types, based on information drawn from other sources or data files, it would be costly and time-consuming, with, we suspect, only marginal improvement in our results.

Price Adjustment Coefficients

Space limitations permit only very sketchy description of the methods used to estimate empirically the effects of inflation on income, asset, and debt types and to derive the corresponding price level elasticities or "adjustment coefficients"; details are given in an extended appendix which will be furnished on request. The coefficients used in the second step (the simulated-inflation) are summarized in Table 1 (for net worth components) and

Table 2 (for income components). An adjustment coefficient is defined as the proportionate change (x) in the money value of an income, asset, or debt type (X) associated with a change in the rate of inflation by amount p . Given a value for x/p and a hypothetical value for p , the proportionate change in the real value of X associated with p is given by $(x - p)/(1 + p)$.

Net Worth Adjustment Coefficients

The coefficients for all claims and debts in the model, whether business or household, have been set equal to zero. Market interest rate adjustment to inflation (partially neutralizing the redistribution effects of the debtor-creditor mechanism) is recorded on income account for those instruments "turning over" within the period. For outstanding marketable instruments, it was decided on the basis of timing considerations to value at face and treat interest flows as fixed in money terms until maturity, rather than to value at market and include in interest flows the annual appreciation in market value as the instrument approaches maturity.

The adjustment coefficients for all tangible assets, whether held by consumer units or firms, are equal to one. Inflation-associated adjustment of market discount rates reflects changes in both the expected rate of inflation and the expected rate of return of tangible assets; the present values on such assets should therefore increase in the same proportion as the price level.

For the types of business equity held by consumer units, adjustment coefficients are based upon the structure of balance sheets constructed for the various business sectors; additions to equity through business saving are accounted for in the income models. With net claims ($NC = \text{claims minus debts}$) fixed in money terms and the market value of tangible assets ris-

ing in proportion to the price level, the adjustment coefficient for business net worth (NW) is $(1 - NC/NW)$.

Income Model Adjustment Coefficients

To test the hypothesis that real wages fall in inflation because of a lag of money wages behind prices, we have carried out (regression) tests based on the behavior during inflation of the share of wages and salaries in total income. We interpret the wage-lag hypothesis as implying an inverse relation between the rate of change (rather than the level) of prices and real wage rates, with real forces (represented by real income or productivity) held constant. The results of separate tests run for the total economy and different sectors, with varying hypothesized lags, failed to support either the wage-lag hypothesis or a contrasting hypothesis of a wage lead. The adjustment coefficient for all wage and salary incomes was therefore assumed to be one.

The interest-income adjustment coefficients are derived from estimates of interest-rate adjustment coefficients and of asset turnover ratios (a reinvestible funds concept). The interest rate coefficients for marketable government securities, corporate and foreign bonds, and mortgages and other loans were estimated through functions including Almon distributed lags of current and past rates of price change on levels of long and short rates.

The coefficient for the commercial bank time deposit rate is based on an estimate of the price-level elasticity of commercial bank net income, derived in turn from estimates of interest-rate adjustment coefficients and turnover ratios for bank assets and on the behavior of bank operating expenses in inflation. The coefficients for the rates paid by other intermediaries were estimated both through this net-income approach and as the product of the elasticity of the commercial bank time deposit

rate with respect to prices and the elasticity of the particular intermediary rate with respect to the (competing) time deposit rate, with quite similar results. The adjustment coefficient for the U.S. savings bond rate was derived through the second technique.

The dividend adjustment coefficient is the product of the elasticity of dividends with respect to corporate (after-tax) cash flow and the elasticity of cash flow with respect to prices. The first elasticity was estimated through a Lintner-type dividend model, involving geometrically decaying lags; the second, through a model containing four mechanisms capable of producing alterations in the real value of corporate net income in inflation: lagged adjustment of labor costs, of net interest costs, and of capital consumption allowances (due to the valuation of depreciable assets at original rather than replacement cost), and incorporation of inventory gains in book profits due to conventional accounting methods.

The adjustment coefficient for retained earnings is based upon coefficients for both dividends and "true" corporate net profits after tax (i.e., exclusive of inventory gains and net of capital consumption based on replacement cost of assets). The latter coefficient is negative since the corporate income tax is levied on both inventory gains and the difference between replacement cost and original cost depreciation. In addition, the dividend coefficient is positive and large since dividends, despite the inertia in payout behavior, are a function of an income concept which includes inventory gains and gains resulting from the difference between replacement and original cost depreciation.

The adjustment coefficients for (true) before-tax profit items—corporate profits before tax (both publicly- and closely-held) and income of unincorporated enterprise—were estimated through models

taking account of possible lagged adjustment of labor and net interest costs. For (true) after-tax income of closely-held corporations, estimates of coefficients for nonfinancial corporations were used. For (true) net rental income, separate adjustment coefficients were estimated for gross rental rates and for landlord costs, the former through a function including Almon distributed lags of current and past rates of price change on the rate of change of a rental rate index, the latter as a weighted average of coefficients for various components of cost.

The pension income adjustment coefficient was calculated as an average of coefficients for the various pension components, using weights derived from OBE control totals; this coefficient varies with p since the speed of Congressional action on Social Security benefits seems to be a function of the rate of price increase itself. Little can be said regarding the inflation sensitivity of the "other income" items (government transfer payments such as public assistance, unemployment compensation, and veteran's benefits, as well as private transfers), and this coefficient has been set equal to zero pending the results of further research.

Effect of the Simulated Inflation on Size Distribution

The results of the simulated inflation, using alternatively a rate of 2% and one of 5%, are summarized in Tables 3 through 7, the first for the net worth model, and subsequent ones for income. Income concept 1 (Table 4) is simply the Census definition of money income used in the SFCC, except that retained earnings of closely-held corporations are included in the incomes of their owners. Income concept 2 (Table 5) adds to money income our imputation of retained earnings of publicly-held corporations to shareholders, in order to place the latter on a

more nearly comparable basis with owners of unincorporated enterprises and recipients of rental income. The effort to put all income recipients on the same pre-direct tax basis is shown in Table 6, where for income concept 3 we have added to concept 2 our imputation of the corporate profits tax to shareholders. Finally, in order to show how the incorporation of real capital gains and losses resulting from inflation would affect the income model, we have added to concept 2 the change in real net worth for each consumer unit as estimated in the net worth model (income concept 4 in Table 7).

In each table consumer units are ranked by size of initial net worth or income and grouped into "quantile" or "fractile" classes of varying size, with much smaller sized groups for the upper tail, in order to take advantage of the larger sample sizes in that part of the distribution. Ranking consumer units by initial net worth or income for the after-inflation distributions has the advantage of showing precisely how those consumer units in the initial quantile fared on the average in the simulated inflation, without any alteration in the composition of the group simply because some fared better or worse than the average. While the results obtained from the more standard Lorenz curve practice of reranking units by their after-inflation net worths or income are also of interest, we found that in all cases but one they differed so little from the unreranked distributions that they were not worth showing separately in the tables. Table 7, however, does give both unreranked and reranked results for income concept 4, where some differences are noticeable.

The tables give two measures of changes in distribution: percent change in the mean amount of real net worth or income for consumer units in, and the percent change in the relative share of, each

quantile. The latter may simply be interpreted as the percent change in the mean for the quantile relative to the mean for the distribution as a whole. Neither of these two measures, of course, reveals differential effects upon those consumer units included in a given quantile.

The two sets of percentage changes differ very little, since the effect of the simulated inflations on the real value of the aggregates themselves is minimal—the largest effect is a 0.7% fall in income concept 2 for the 5% inflation. Our discussion is therefore confined to the relative share or relative mean income measure.¹ Contrary to expectation, aggregate real net worth actually rises, though by a miniscule amount (0.06% for the 5% case). This latter result is undoubtedly attributable to reporting deficiencies in the file: claims are underreported by consumer units by a greater proportion than either stockholdings or debts.

The results for net worth show a small but fairly consistent shift to less inequality, with the shares of quantiles below the 80th percentile rising, and those above, falling, although the very top groups appear to lose a bit less than those immediately beneath them.

These results can readily be accounted for by the fact that the ratio of net claims to net worth at the bottom of the distribution is large and negative, falling in absolute value until it becomes positive at about the 80th percentile, with little change

¹ If g is the ratio of mean net worth or income for the entire distribution before and after the simulated inflation, h , the corresponding ratio for a given quantile, and k , the ratio of the relative mean incomes or shares before and after inflation for the same quantile, then $k = h/g$. Since in our simulations g turned out to have a value close to unity and h had a value "close" to k , we have confined our discussion in the text to $(k-1)$, although the values for both $(h-1)$ and $(k-1)$, expressed as percents, are shown in the tables. Even though "small," $(h-1)$ and $(k-1)$ can differ in sign if $g < h < 1$ or $g > h > 1$.

TABLE 3—EFFECT OF SIMULATED INFLATION ON VALUE OF AND SHARES IN REAL NET WORTH FOR QUANTILE GROUPS*

Size of Quantile	Mean Net Worth of Quantile	Relative Share	Pct. Change in Shares		Pct. Change in Real Value	
			$p=0.05$	$p=0.02$	$p=0.05$	$p=0.02$
1-10	\$-415	-0.214	10.546	4.386	10.485	4.340
11-30	248	0.247	11.048	4.614	11.099	4.593
31-40	2,000	0.997	5.026	2.087	5.084	2.104
41-50	4,503	2.245	3.688	1.528	3.742	1.548
51-60	7,479	3.729	1.523	-0.631	1.578	0.653
61-70	11,328	5.648	0.754	0.312	0.809	0.335
71-80	16,901	8.427	0.189	0.079	0.245	0.101
81-90	27,767	13.845	-0.406	-0.168	-0.349	-0.145
91-95	47,150	11.755	-0.413	-0.170	-0.356	-0.148
96-100	213,900	53.530	-0.361	-0.150	-0.303	-0.128
top 1%	\$662,183	33.016	-0.312	-0.131	-0.254	-0.109
top $\frac{1}{2}$ %	1,565,799	19.517	-0.237	-0.100	-0.179	-0.078
all units	\$20,050	100	—	—	0.058	0.022

* Ranked by size of initial net worth

thereafter. This is primarily because the debt/asset ratio falls drastically with increases in net worth; the share of claims in total assets appears to be relatively constant, although there is a marked shift from liquid assets to marketable claims as net worth increases. The mitigating effect at the top arises from the very large share of common stock held by the top groups compared with those immediately below them, who hold relatively more assets in real estate (for which our adjustment coefficients are less than for stock).²

Perhaps the most striking result for the income models is the relatively small effect our simulated inflations of either the 2% or 5% form have on the relative distribution. Furthermore, the effect of the 2% case is generally less than half the 5% case, except for the bottom four deciles, where pension income is concentrated and hence where the difference in assumptions concerning the adjustment coefficients for

such income would be expected to have its primary impact.

What change there is seems to be of a rather consistent and plausible type. For money income (concept #1), the before and after Lorenz curves intersect. The lower groups and the upper tail appear to lose relative to the middle in terms of shares (although the mean income of all quantiles is reduced). For the 5% case the relative share (and hence the relative mean income) of the bottom two quintiles and of the top 3% are reduced by 0.7% and $\frac{1}{2}$ % respectively, whereas the share of the quantile composing the 41st through the 96th percentiles is increased by $\frac{1}{4}$ %.³ (The corresponding figures for the 2% case are -0.3, -0.2, and +0.1 respectively.)

These changes are more or less what would be expected. The middle quantile previously referred to gets most of its income from earnings (wages and salaries and self-employment income), for which

² Information on the composition of net worth by size of net worth may be found in [3, Tables A 8-A 14].

³ For a discussion of the method for pinpointing these changes and for their interpretation, see [1].

TABLE 4—EFFECT OF SIMULATED INFLATION ON VALUE OF AND SHARES IN REAL MONEY INCOME (INCOME CONCEPT #1) FOR QUANTILE GROUPS*

Size of Quantile	Mean Income of Quantile	Relative Shares	Pct. Change in Shares		Pct. Change in Real Value	
			$p=0.05$	$p=0.02$	$p=0.05$	$p=0.02$
1-10	\$ 533	0.667	-1.604	-0.824	-2.017	-1.017
11-20	1,568	2.453	-1.114	-0.603	-1.541	-0.793
21-30	2,603	4.075	-0.679	-0.342	-1.088	-0.532
31-40	3,624	5.671	-0.169	-0.107	-0.580	-0.298
41-50	4,664	7.300	0.145	0.066	-0.267	-0.124
51-60	5,778	9.058	0.239	0.098	-0.173	-0.093
61-70	6,968	10.905	0.282	0.132	-0.130	-0.059
71-80	8,289	12.972	0.232	0.106	-0.180	-0.084
81-90	10,249	16.039	0.230	0.109	-0.182	-0.081
91-95	13,076	10.232	0.221	0.105	-0.165	-0.077
96-100	26,384	20.645	-0.425	-0.206	-0.789	-0.352
top 2%	\$40,949	12.816	-0.807	-0.378	-1.179	-0.522
top 1%	57,943	9.067	-1.027	-0.428	-1.399	-0.604
top $\frac{1}{2}$ %	78,178	6.117	-1.292	-0.592	-1.665	-0.764
top $\frac{1}{4}$ %	105,461	4.126	-1.478	-0.614	-1.884	-0.803
all units	\$6,391	100	—	—	-0.413	-0.192

* Ranked by size of initial income #1

our adjustment coefficients are close to one. While the majority of the income of the units in the bottom 40% is from earnings too, they are also dependent on transfer payments, while the top groups are more heavily dependent on property income.⁴

Our simulated inflation reduces the share of the top income groups even more when undistributed profits are allocated to shareholders (income concept #2), intensifying somewhat the Lorenz curve intersection phenomenon. Using the three quantile groupings just defined and the 5% inflation case, we find that the share of the bottom two quintiles fell by 0.4%, that of the middle group rose by 0.5%, and that of the top 3% fell by 1.5%. As noted before, retained earnings in our model have a negative adjustment coefficient.

⁴Information on the composition of income by size of income (SFCC concept) may be found in [4, Tables 1-3].

That these Lorenz curve intersection phenomena are mostly the result of including in the income of shareholders only part of the pretax income of corporations can be seen more clearly when an attempt is made to allocate all corporate profits before tax to the income of shareholders (income concept #3, Table 6). In this case the after-inflation distributions show an observable, though quite small, movement towards more inequality. In the 5% case, for example, the lower half of the distribution loses relative to the upper half; the share of the former falls by 0.5%, that of the latter rises by 0.1%. It is interesting to note that two of the more important redistributive forces at work lie within the domain of government tax and transfer policy. A distribution based on incomes after direct taxes, both income and corporate, might well show more of a reduction in the share of the upper groups than does income concept #2.

While the impact of inflation appears to

AMERICAN ECONOMIC ASSOCIATION

TABLE 5—EFFECT OF SIMULATED INFLATION ON VALUE OF AND SHARES IN REAL INCOME INCLUSIVE OF UNDISTRIBUTED PROFITS (INCOME CONCEPT #2)*

Size of Quantile	Mean Income of Quantile	Relative Shares	Pct. Change in Shares		Pct. Change in Real Value	
			$p=0.05$	$p=0.02$	$p=0.05$	$p=0.02$
1-10	\$ 544	0.664	-1.504	-0.782	-2.210	-1.099
11-20	1,578	2.412	-0.914	-0.510	-1.633	-0.828
21-30	2,617	4.001	-0.388	-0.221	-1.111	-0.539
31-40	3,651	5.581	-0.057	-0.013	-0.669	-0.333
41-50	4,691	7.170	0.435	0.184	-0.293	-0.135
51-60	5,830	8.911	0.494	0.204	-0.234	-0.115
61-70	7,017	10.727	0.476	0.210	-0.252	-0.110
71-80	8,329	12.731	0.515	0.222	-0.213	-0.098
81-90	10,356	15.830	0.433	0.194	-0.295	-0.128
91-95	13,319	10.180	0.371	0.168	-0.356	-0.152
96-100	28,533	21.807	-1.192	-0.552	-1.872	-0.873
top 4%	\$31,768	19.423	-1.387	-0.603	-2.067	-0.976
top 1%	65,890	10.072	-2.317	-0.954	-2.992	-1.257
top $\frac{1}{2}$ %	90,823	6.941	-2.786	-1.300	-3.455	-1.502
top $\frac{1}{4}$ %	126,494	4.834	-3.337	-1.366	-4.038	-1.680
all units	\$6,543	100	—	—	-0.726	-0.320

* Ranked by size of initial income #2

TABLE 6—EFFECT OF SIMULATED INFLATION ON VALUE OF AND SHARES IN REAL INCOME INCLUSIVE OF CORPORATE PROFITS BEFORE TAX (INCOME CONCEPT #3)*

Size of Quantile	Mean Income of Quantile	Relative Shares	Pct. Change in Shares		Pct. Change in Real Value	
			$p=0.05$	$p=0.02$	$p=0.05$	$p=0.02$
1-10	\$ 553	0.647	-1.807	-0.911	-2.010	-1.011
11-20	1,592	2.331	-1.279	-0.666	-1.488	-0.769
21-30	2,632	3.854	-0.815	-0.396	-1.025	-0.499
31-40	3,669	5.372	-0.340	-0.178	-0.551	-0.282
41-50	4,722	6.914	-0.009	-0.003	-0.222	-0.106
51-60	5,864	8.586	0.056	0.021	-0.156	-0.072
61-70	7,071	10.353	0.073	0.041	-0.139	-0.063
71-80	8,395	12.292	0.054	0.016	-0.188	-0.088
81-90	10,428	15.269	0.071	0.040	-0.141	-0.064
91-95	13,601	9.957	0.123	0.060	-0.089	-0.043
96-100	33,378	24.495	0.224	0.102	-0.018	-0.008
top 2%	\$56,262	16.511	0.233	0.113	0.019	0.009
top 1%	83,304	12.227	0.241	0.118	0.021	0.015
top $\frac{1}{2}$ %	120,680	8.835	0.337	0.166	0.125	0.063
top $\frac{1}{4}$ %	174,737	6.396	0.351	0.179	0.137	0.075
all units	\$6,830	100	—	—	-0.213	-0.104

* Ranked by size of initial income #3.

TABLE 7—EFFECT OF SIMULATED INFLATION* ON VALUE OF AND SHARES IN REAL INCOME INCLUSIVE OF UNDISTRIBUTED PROFITS AND INFLATION-INDUCED REAL CAPITAL GAINS AND LOSSES (INCOME CONCEPT #4)

Size of Quantile	Mean Income of Quantile	Relative Shares	Pct. Change in Shares		Pct. Change in Real Value	
			Unreranked ^b	Reranked ^c	Unreranked ^b	Reranked ^c
1-10	\$ 544	0.664	-7.845	-8.371	-9.702	-10.213
11-20	1,578	2.412	-5.343	-5.427	-5.884	-5.968
21-30	2,617	4.001	-1.468	-2.706	-2.030	-3.261
31-40	3,651	5.581	-0.738	-0.924	-1.305	-1.489
41-50	4,691	7.170	-0.396	-0.128	-0.176	-0.697
51-60	5,830	8.911	1.592	0.889	1.012	0.314
61-70	7,017	10.727	0.827	1.038	0.251	0.462
71-80	8,329	12.731	0.642	1.288	0.067	0.710
81-90	10,249	15.830	1.556	1.689	0.976	1.109
91-95	13,319	10.180	0.082	0.089	-0.489	0.316
96-100	28,533	21.807	-1.278	-0.915	-1.804	-1.441
top 3%	\$31,768	16.850	-1.783	-1.365	-2.333	-1.914
top 1%	65,890	10.072	-2.555	-2.094	-3.111	-2.653
top ½%	90,823	6.941	-3.925	-2.297	-4.473	-2.854
top ¼%	126,494	4.834	-5.154	-2.792	-5.695	-3.345
all units	\$6,543	100	—	—	-0.571	-0.244

* Only the $p=0.05$ case is presented.

^b Ranked by size of initial income #4.

^c Ranked by size of after-inflation income #4.

be small for both net worth and income concepts, the patterns of change differ; the net worth model shows reduced inequality, whereas income concepts 1 and 2 show reduced inequality in the upper part of the distribution and increased inequality in the lower part, and income concept 3 shows an increase in inequality throughout. Space is lacking to analyze the sources of the differences, but a few factors which should be taken into account in their reconciliation may be pointed out. Certain money income items do not have counterparts in the consumer's balance sheet, at least as conventionally measured. In particular, human capital and the value of claims to future transfer payments are omitted, making the wealth distribution more nearly relevant to the distribution of property income than of total income. On the other hand, certain components of net worth have no correspondent (negative or

positive) income counterpart. Unsecured consumer debt is deducted from assets in computing net worth, but interest paid on such debt is not deducted from interest received by consumer units, either in field survey income concepts or in the concept of personal income. No money income type corresponds to owner-occupied dwellings and checking accounts, and only money interest received, to savings accounts, in the money income concept, although imputations for such missing rent and interest are included in OBE's personal income. The omission of wealth items, for example, is probably sufficient to account for the difference between our findings for net worth and the Lorenz curve intersection phenomenon for the first two income concepts.

One way of trying to increase the comparability of income and wealth concepts is by combining them into one measure.

Out of the many candidates we have experimented with only one here: the inclusion of inflation-induced real capital gains and losses in income inclusive of corporate saving. Table 7, which summarizes the results, omits the 2% inflation case in order to show the effect of reranking consumer units by the size of their after-inflation income, the only concept for which reranking makes any real difference. Whichever way it is done, however, the bottom group and the upper tail appear to lose more to the middle than is true for the first two concepts. Unreranked, the shares of the bottom two quintiles and the top 3% are reduced by 2.5% and 1.8% respectively and that of the middle group is increased by 0.9%. The greater negative effect at the bottom is due to the fact that net indebtedness is low and assets are mostly in the form of claims, particularly for the aged units which make up much of this group. In the middle group, on the other hand, indebtedness is larger. The top groups have little debt; in addition, the negative effect arising from their holdings of claims more than offsets the positive wealth effect from stockholdings.

In summary, while much more could be said on the forms and extent of redistribution revealed by this study and more re-

finements could be introduced into the models themselves, they are not likely to alter our findings on the rather modest extent of the redistributive effects of changes in the rate of inflation or on the patterns of change in the distribution of real income (variously defined) or net worth that the simulations reveal. Further, if the rather mild effects noted here for income are deemed to be undesirable ones, they can be mitigated by modifications of existing transfer and tax policies.

REFERENCES

1. Edward C. Budd, "Postwar Changes in the Size Distribution of Income in the U.S.," *A.E.R.*, May 1970, pp. 247-60.
2. Robinson G. Hollister and John L. Palmer, "The Impact of Inflation on the Poor," Discussion Paper of the Institute for Research on Poverty, University of Wisconsin, Madison, 1969.
3. Dorothy S. Projector and Gertrude Weiss, *Survey of Financial Characteristics of Consumers* (Board of Governors of the Federal Reserve Systems, 1966).
4. Dorothy S. Projector, Gertrude S. Weiss, and Erling T. Thoresen, "Composition of Income as Shown by the Survey of Financial Characteristics of Consumers," *Six Papers on the Size Distribution of Wealth and Income*, Lee Soltow, editor (National Bureau of Economic Research, 1969).

The Two-way Relationship between the Budget and Economic Variables

By ARNOLD H. PACKER

*U. S. Office of Management and Budget**

In theory, the level and composition of the Federal budget are set to meet the policymaker's objectives for stabilization, allocation, distribution, and efficiency. Unfortunately, the analytical framework that exists supports stabilization objectives—and these only in the short run.

The aim of this paper is to investigate the effect of a hypothetical major budget decision on both stabilization and allocation variables over a five-year planning horizon.

First, the endogenous nature of many budget outlays are recognized by defining relationships between outlays and general macroeconomic variables. Second, a framework is established for estimating the net allocative effect of the budget. Finally, the paper uses a simple macroeconomic model to trace out the effects of a choice between two alternative budgets on both stabilization and allocation and thus attempts to synthesize these aspects of budget policy. The methodological approach used is not one which provides an estimate of the "true" relationships; rather it draws out the implications of a set of "reasonable" assumptions.

The analytical picture may be made more concrete by considering the central problem of economic policy as it appeared in the spring and summer of 1970 when the planning cycle for fiscal year 1972 was initiated. A moderate (4%) gap had already grown between potential and actual

output, and the real growth rate was still less than potential so that the gap was growing. Forecasts indicated that, given the policy actions already taken, the gap would be approximately 5%, and its rate of change would be approximately zero by the start of fiscal year 1972. The Administration objectives were to close the gap as quickly as possible without recreating an inflationary situation and to allow a recovery in housing and the other sectors that had borne the burdens of restrictive monetary policy. The chosen economic policy would help determine the answers to the following questions:

1. The stabilization question: How quickly would the gap be closed?
2. The allocation question: What types of economic activity would fill the gap?

The policy answers were to be reflected in the FY 1972 budget targets and were made in the context of a five-year planning horizon. The analytical task—as is generally the case when the objective function is not stated explicitly—was to estimate the alternative outcomes corresponding to alternative decisions. In this case, it was alternative stabilization and allocation outcomes for alternative feasible FY 1972–76 streams of budget outlays. In this analysis, contrary to the traditional approaches, outlays were not to be considered completely exogenous. That is, a more realistic analytical framework was to be used, one in which only a small portion of budget outlays is controlled in any budget year.

* The opinions expressed are those of the author and do not necessarily represent the views of the Office of Management and Budget.

I. Stabilization

A. *The Impact of the Budget on Economic Activity*

Almost all economic models—from the simplest textbook expositions to the most complete econometric models—explicitly recognize the macroeconomic relationships running from government expenditures to economic activity and from the economy—via tax rates—to government receipts.

The net impact of the budget on economic activity has been quantitatively estimated by many methods. One method is the use of increasingly comprehensive economic models that contain federal purchases and some, or all, of other federal expenditures as exogenous variables. Another method attempts to reduce the fiscal impact to a single number such as the full employment surplus (FES). Sophisticated applications of both methods recognize that every dollar of federal outlays will not have the same economic impact.

A third method—the Reduced Form—has recently been suggested by Hymans and Wernette.¹ They define fiscal stimulus as the measured year-to-year increase in expenditures plus the year-to-year decrease in taxes due to changes in tax rates (but excluding any induced changes in revenues). Conceptually, the FES and Reduced Form methods eliminate induced, nondiscretionary results of discretionary budget decisions. Empirically they do not, because they ignore all (or almost all) of the induced changes in expenditures.

It may be futile to try to determine if a particular budget is absolutely stimulative, or restrictive, without relating the budget to either price/unemployment goals or some long-term trend of growth in output. What is necessary is an economic forecast that includes a forecast of the endogenous parts of the budget, in both the receipt and expenditure accounts.

¹ "The Impact of the Federal Budget on Total Spending," *Business Economics*, September 1970, pp. 29–34.

B. *Response of Budget Outlays to Changes in Macroeconomic Variables*

This section defines the relationships between budget outlays and the macroeconomic variables. The relationships will provide estimates of "nondiscretionary" expenditures for the 1972–76 planning horizon.

"Nondiscretionary" is rather difficult to define in this context. At one extreme, one can ignore all political realities and assume that only existing laws and contracts limit the policymaker's discretion. At the other extreme, one can assume that the policymakers cannot diminish the growth of expenditures for health, education, etc., but can only shift expenditures within relatively narrow categories. Some compromise between these two extremes is necessary.

Two bodies of expenditure data are available. One is the data of past rates of growth; these commingle discretionary and nondiscretionary effects and tend to understate the room for discretionary action. The second body of data has been developed in the Office of Management and Budget over the last two years. These are program-by-program forecasts that include individual judgments as to administration commitments, political necessity, etc. These data tend to overstate the room for discretionary action.

The relationships described below draw on both bodies of data.² They are judgments of "reasonable relationships" and not testable hypotheses.

Government-wide pay raises are granted periodically in order to make government wages and salaries "comparable" to those in the private sector. The new law makes the increases automatic beginning in 1972. The comparable private sector pay in-

² I wish to thank Charles Mohan, Darwin Johnson, Maureen Smet, Douglas Norwood, and others for their help with much of the quantitative work used to develop these relationships.

crease, measured in May of each year, will be included in a pay raise that will take effect in October of the same year.

Assuming private sector real wages will increase with output per man and a productivity increase of 3.2% annually, this arrangement can be approximately represented by:

$$(1) \text{ Comp}_T = \text{Comp}_{T-1} (1 + .032 + \Delta \text{GNPD}_{t-1})$$

where the upper case T refers to fiscal years and the lower case t to calendar years and GNPD is the GNP deflator.³

The increase in prices also affects non-compensation purchases (primarily goods). Budget policy has generally been to use prices existing at the same time the budget is prepared to cost out next year's programs. Therefore, one-half of the price increase taking place in $FY (1)$ —i.e., the July to January price increase—will affect $FY (2)$ budget decisions; all of the increase will affect future budgets.⁴

$$(2) \text{ Goods Deflator}_T = \text{Goods Deflator}_{T-1} (1 + \Delta \text{GNPD}_{T-1})$$

Defense purchases are, relatively, the most controllable of all Federal outlays. The size of the Armed Forces, the number of civilians working for the Department of Defense, and the real level of defense goods purchases are, within limits, subject to the discretion of the policymakers.

The exercise that will be used to illustrate the ideas contained in this paper will

³ Regressing annual 1952-69 data gives the following results:

$$\begin{aligned} \log (\text{Comp})_t &= 8.155 + 1.18 \log (\text{GNPD})_{t-1} \\ &\quad (533) \quad (2.5) \\ &\quad + 1.08 \log (\text{Private Productivity}) \\ &\quad (3.5) \end{aligned}$$

(The t -tests are in parentheses below the coefficients.)

⁴ The real value of goods bought in $FY (1)$ and (2) will, of course, be diminished to the extent that prices increase.

examine a hypothetical choice between a "high" and a "low," five-year defense budget. The "high" defense budget will maintain the fiscal 1971 level of military and civilian defense manpower throughout $FY 1972-76$, it will also increase the real level of defense spending for goods by 5% in $FY 1972$ and then reduce that level of real spending by $2\frac{1}{2}\%$ each succeeding year through $FY 1976$.

The "low" defense budget will decrease defense manpower by 10% in $FY 1972$ and by approximately 4% each succeeding year; also, it will decrease real defense goods purchases by 10% in $FY 1972$ and by approximately 3% each succeeding years. The price the government will have to pay for these real resources will be determined, in part, by the effect of the decision (between the "high" and "low" defense budgets) on the general price level.⁵ All of the remaining budget outlays will be endogenously determined.

Non-defense purchases will amount to \$24 billion in fiscal 1971. Compensation will account for approximately 45% of these purchases. The exercise will assume that the real level of purchases of goods will remain constant throughout 1971-76, the price index being given by Eq. 2.⁵

Compensation per man is given by Eq. (1). The number of employees of the non-defense agencies is the result of two offsetting trends: the increasing productivity of government workers and the increasing workload of the nondefense agencies. Productivity is implicitly recognized in the budget process by the requirement that the agencies "absorb" a portion of each government-wide pay increase.

The nondefense workload increase is assumed to be equal to the increase in potential GNP; productivity increases are assumed equal to that experienced by the

⁵ This assumes that the budget requests are for a higher level of real purchases, but that inflation, in each year, prevents the intended increase.

private sector. Thus, the built-in growth in nondefense Federal employment is approximately equal to the private sector increase in man-hours, or about 1% annually.

Grant programs are frequently uncontrollable from the point of view of the federal decision-maker once the legislation has been enacted.

Certain grant programs depend on prices in a manner similar to nondefense purchases of goods. This is true for any construction program where the accepted objective is measured in real terms; for example, the cost of completing the interstate highway system is one of the determinants of highway grants.

The historical data indicate that price increases are significant variables in predicting grant outlays though the coefficient is significantly more than one (even when a trend variable is included). The assumption used in the model allows grants to increase in step with prices and potential GNP. Thus,

$$(3) \text{ Grants}_T = \text{Grants}_{T-1} (1 + .043 + \Delta \text{GNPD}_{t-1})$$

Transfers.—Outlays for many transfer programs depend, by law, on macroeconomic variables. Approximately 65% of transfer outlays in fiscal 1971 are for retirement and social insurance programs. If one version of the legislation currently being considered by Congress becomes law, then future social security benefits will be tied to the consumer price index in the following way. Beginning in 1972, benefit levels will be increased in January to compensate for the increase in the CPI for the 12 months ending in the previous May (if the last 12-month CPI increase exceeded 3%).

This legislation appears to affirm the view that Social Security benefit recipients are entitled only to preservation of the real purchasing power of their benefits, but not to parity with current wages. That is,

beneficiaries do not benefit from post-retirement productivity increases. The data indicate that, in the 1953-69 period, transfer payment recipients obtained a compromise between these two positions (i.e., between prices vs prices plus productivity). It is unlikely that future increases in transfers will be limited to price increases.

Unemployment compensation accounts for about 5% of transfer programs. Outlays for this purpose now increase by about \$125 million for each maintained .1% increase in the global unemployment rate.

Assuming 2% benefit enrichment in the future (i.e., 2/3 of productivity) and 1% increases in the beneficiary population annually;

$$(4) \text{ Transfer}_T = \text{Transfer}_{T-1} (1 + .03 + \Delta \text{GNPD}_{t-1}) + .125(U_T - U_{T-1})$$

Subsidies.—The determinants of "subsidies less current surplus of government enterprises" have not received much attention. Much of the short-term variation can be explained by postal pay increases and postal rate increases. The pay component can be treated as the other pay items (see equation (1)).

Interest.—Net interest paid on the debt is the most endogenous of Federal outlays. Changes in interest payments depend on the turnover of existing debt (and the average interest rate on that debt), the average interest rate at which new debt must be financed, and the change in the size of the debt. Only half the full year interest need be paid on the deficit that is accumulated uniformly during a fiscal year. This is approximated by using the previous calendar year ($t-1$) deficit to compute the interest payment for fiscal year T . Future interest rates are assumed equal to a real interest rate of 3.5% plus the inflation rate.

$$\begin{aligned}
 (5) \quad \Delta \text{Interest payments}_T & \\
 &= (.035 + \Delta \text{GNPD}_{t-1}) \text{Deficit}_{t-1} \\
 &\quad + (.035 + \Delta \text{GNPD}_{t-1} \\
 &\quad \quad - \text{old rate}) \text{Turnover}_{t-1}
 \end{aligned}$$

The turnover is assumed to be \$100 billion annually and the old rates, for FY 1972-76, are approximately 6¼%.

It should be clear that alternative budgets in one year—because they lead to alternative price levels, unemployment rates, deficits, and interest rates—lead to different budget situations in succeeding years.

Some changes, e.g., unemployment, have only a temporary effect on budget outlays; others, such as the size of the deficit or price changes have a once-and-for-all effect on the budget level. The final section of the paper will trace out some of these effects under alternative budgets. Before that, however, a framework will be established for examining the allocation problem.

II. Allocation

Public finance texts describe the federal budget's allocative role as the allocation of resources between federal and other (private, and state and local) uses and the allocation of the federally-employed resources. The allocative effect is greater than this, however. For example, the size of the deficit itself, as well as federal credit programs, affect the interest rate, and thus, the allocation between consumption and investment.

Since the policymakers cannot avoid affecting the allocation of resources, it is worthwhile to be able to estimate the allocation of GNP (among health, education, transportation, etc.) that will follow from a specified set of policy actions. Therefore, it is appropriate to study the effect of the federal budget on aggregate private allocation decisions.

In addition, if increasing individual and local government discretion in allocation decisions is an objective, then it is important to know which sector will be making the purchases. That is, the allocation analysis should seek to answer the following questions for a number of functions (health, education, transportation, housing, general government).

1. What share of GNP will be devoted to this function?
2. How much of the share will be public? How much of the private share will be indirectly supported by government grants or transfers?
3. How much of the public share will be federal? How much of the State and local share will be indirectly supported by Federal grants?

An analysis of this sort can lead to normative questions concerning functional GNP shares. What is the desirable course of real per capita expenditures on health or on education? Is this a meaningful way in which to state the question for defense? If not, what objective considerations are meaningful in determining the appropriate share of GNP (or of the budget) for defense? Which functional shares of the budget and/or GNP shall diminish to accommodate the growing shares?

Ultimately, the analysis could be extended so that alternative answers to the above questions could be evaluated in terms of specific objectives. For example, will an increase in the public share of transportation (e.g., mass transit) have a greater impact on health objectives by reducing accidents and pollution than a corresponding increase in the public share of health (e.g., some form of national health insurance)?

The framework seems necessary if we are to have a way of analyzing the impact of a national health insurance system, or appraising ideas such as the proposal that

the federal government assume the full burden of the welfare system in exchange for its share in supporting the education system.

The allocation analysis was made in the following manner. First, nine exclusive and exhaustive functional categories of GNP were defined: basic necessities (food, clothing and rent), health, education, transportation, defense, general government, new housing, business investment and all others.

Second, long run trends of GNP shares for each function were projected to 1976.⁶ Relationships between implicit deflators for each function and the GNP deflator were also defined so that the allocation of real resources could be computed. Finally, relationships were defined between changes in fiscal and monetary policy variables and deviations of the GNP shares from their long term trends. No claim is made for the validity of the quantitative relationships. Reasonableness rather than rigor was the criterion used. The objective was to get reasonable estimates of the differences in allocation that would follow from the alternative budgets.

The final section illustrates these concepts.

III. *An Analysis of the Implications of Alternative Budgets*

Postulate the following simplified decision-making and economic environment. The fiscal decision maker can decide between the two five-year defense budgets described earlier; the higher defense budget is \$10-18 billion more than the lower for each year of his five-year planning horizon. In general, he wants to balance

the full employment budget. However, if the sum of the defense and built-in domestic program outlays exceeds full-employment revenues and unemployment exceeds 4%, the policymaker will tolerate a full employment deficit. He has ranked his domestic budget priorities in the following way. The first billion dollars that becomes available will go to education, the second to general government (say, for pollution abatement), the third to health, the fourth to transportation; and then, if more funds are available, the cycle will be repeated. Once a decision is made to raise the spending level for some purpose, the higher level will continue for the rest of the planning horizon. All the new programs are grant-in-aid programs.

There is an independent monetary authority who shares the fiscal policymaker's economic model (but not necessarily his price/employment objectives). The monetary authority uses a quadratic function of last year's unemployment and inflation performance to determine this year's increase in the money supply; it is assumed to be:

$$M_t = 1.04M_{t-1} + 200(U_{t-1} - .04)^2 - 10(\Delta P_{t-1} - .025)^2$$

The economic model used by the fiscal and monetary policymaker is described below.

Nominal GNP, Y , will increase by 7% annually if fiscal and monetary policy are neutral; i.e.,:

$$Y_t = 1.07Y_{t-1} + \text{Fiscal Stimulus}_t + \text{Monetary Stimulus}_t$$

Fiscal stimulus in year t is equal to a fraction of the previous year's deficit D plus the sum of this fiscal year's increase in federal purchases and grants (a neutral fiscal policy implies a fiscal stimulus equal to zero). Thus, the multiplier of purchases and grants are assumed to be one (with an average six-month lag) and any second

⁶ A relatively simple regression was used for each function. For example, consumption, the change in consumption, and Federal grants for transportation were used as explanatory variables for transportation; nonmedical transfer payments were used for expenditures for basic necessities.

round effects of a deficit operate with a full year's lag.

Monetary stimulus is determined by the central bank and the government's net fiscal position. If the budget is balanced, then a 4% increase in the money stock represents a neutral monetary position. If the budget is in deficit then an addition (above 4%) to the money supply equal to a fraction of the deficit must be supplied to keep monetary policy neutral. Monetary stimulus is a multiple of the money supplied in excess of the 4% adjusted for the Federal deficit.

The assumed economic model is

$$(6) \quad Y_t = 1.07 Y_{t-1} + .6 D_{t-1} + \Delta \text{Purchases}_t \\ + \Delta \text{Grants}_t \\ + 2.3(M_t - 1.04 M_{t-1} - .15 D_{t-1})$$

The unemployment rate, U , is assumed to be given by a simple representation of Okun's law:

$$(7) \quad U_t = 0.04 + .333(\text{Gap})_t$$

The rate of price increase is assumed to be given by a simple Phillips relationship which includes a price expectation variable:

$$(8) \quad \Delta P_t = .7(\Delta P_{t-1} - .025) + \frac{.0012}{U_t}$$

The year-to-year change in the federal full employment tax yield is assumed to be 20% of the year-to-year difference in full employment GNP (as modified by the 1969 Tax Reform and Social Security Acts). The within-year marginal Federal tax yield is assumed to be 30% of the difference between full employment and actual GNP.

Equilibrium is *not* the initial condition; there is a gap of 4%, unemployment is 5.3% ($4.0 + 4/3$); and the rate of price increase still shows the effects of a previous inflationary period. One simulation was

run for each budget assumption.

The stabilization results are shown in Tables 1 and 2. Tables 1A and 1B show the values for the economic and budget variables, respectively, for the five-year planning horizon under the high-defense budget; Tables 2A and 2B show the variable values associated with the low budget.

If the high-defense budget is chosen, then the sum of defense and "built-in" domestic outlays exceed full employment revenues through FY 1974. If the low-defense budget is chosen, then outlays do not exceed full employment revenues in any year and new initiatives can be introduced in fiscal 1974 and increased to \$18 billion by 1975 (actual deficits are incurred in fiscal 1972 and 1973).

Only in the beginning of the period do the alternative budgets lead to different stabilization results. Choosing the lower budget leads to slightly higher unemployment rates in 1971 and 1972 and slightly lower inflation rates in 1971, 1972 and 1973. By 1973 the gap between actual and potential GNP is approximately 2% irrespective of the budget chosen. The higher unemployment in 1971, associated with the lower budget, leads to more monetary stimulus in 1972. The monetary and budgetary conditions associated with the two budgets lead to a different allocation of resources in each of the 1971-1975 years.

The allocation computation displaces the GNP shares from their long-term trend (1969 to 1975 interpolation). It assumes that monetary stimulus affects the economy through the mortgage and state and local bond markets in a three-to-one ratio. Thus, the change in GNP brought about by any monetary stimulus or restraint is distributed in the following way: 75% to housing; 10% to education; 5% to health; 5% to general government and 5% to transportation.

The additional grants (i.e., the initiatives) are also assumed to increase state

TABLE 1—STABILIZATION RESULTS WITH HIGH DEFENSE BUDGET

Table 1A—Economic Variables by Calendar Year (\$ Billions or %)						
	1970	1971	1972	1973	1974	1975
GNP	979	1070	1168	1269	1376	1500
Deflator	135	140	145	150	155	160
Change (%)	5.0	4.1	3.6	3.4	3.3	3.5
Unemplmt (%)	5.3	5.1	4.8	4.6	4.4	4.1
Real GNP	727	764	805	846	888	936
Potential	758	791	825	860	897	936
Gap (%)	4	3	2	2	1	0
Deficit	-8	-17	-21	-17	-7	-1
Budget Outlays	204	226	250	264	276	297
Revenue	196	210	228	245	267	297
Money Stock	205	219	232	245	256	267
Change (%)	4.0	6.8	6.2	5.3	4.7	4.3

Table 1B—Budget Variables by Fiscal Year (\$ Billions)

	1971	1972	1973	1974	1975	1976
Defense	72	81	87	89	92	95
(Comp)	30	35	40	42	45	48
(Goods)	42	46	47	47	47	47
Non-Defense	25	25	26	27	29	30
(Comp)	11	11	12	13	14	15
(Goods)	14	14	14	14	14	15
Transfers	66	81	85	90	95	101
Grants	28	30	32	34	36	54
(Initiatives)	0	0	0	0	0	16
Subsidies	6	6	6	6	6	6
Interest	16	19	21	23	24	25
Total Outlays	210	242	257	270	282	311
Receipts	198	220	237	256	282	308
Surplus	-12	-23	-20	-14	1	-3
F.E. Rcpts.	212	230	244	259	283	312

and local spending for the relevant functions (education, general government, health and transportation) on a 3-to-2 basis (i.e. an effective 2-to-1 matching basis). In this way both fiscal and monetary policy tend to have an effect on allocation.

Investment is assumed to be more cyclically volatile than the other sectors, and is therefore assumed to vary from its trend percentage of GNP by 10% of the deviation of GNP from its 7% growth trend.

Tables 3A and 4A show the percent distribution of GNP among the nine func-

TABLE 2—STABILIZATION RESULTS WITH LOW DEFENSE BUDGET

Table 2A—Economic Variables by Calendar Year (\$ Billions or %)						
	1970	1971	1972	1973	1974	1975
GNP	979	1062	1164	1266	1379	1499
Deflator	135	140	145	150	155	160
Change	5.0	4.0	3.5	3.3	3.3	3.5
Unemplmt	5.3	5.3	4.9	4.6	4.3	4.1
Real GNP	727	759	804	847	892	937
Pot. GNP	758	791	825	860	897	936
Gap %	4	4	3	2	1	—0
Deficit	—8	—13	—10	—4	0	—1
Bud. Outlays	204	221	238	251	270	296
Revenue	196	208	228	245	268	297
Money Stock	205	219	235	248	260	270
Change	4.0	6.8	7.3	5.5	4.7	4.1

Table 2B—Budget Variables by Fiscal Year (\$ Billions)						
	1971	1972	1973	1974	1975	1976
Defense	72	71	75	76	76	77
(Comp)	30	32	35	36	37	37
(Goods)	42	40	40	40	40	40
Non-Defense	25	25	26	27	29	30
(Comp)	11	11	12	13	14	15
(Goods)	14	14	14	14	14	15
Transfers	66	81	85	90	95	101
Grants	28	30	32	38	54	75
(Initiatives)	0	0	0	4	18	37
Subsidies	6	6	6	6	6	6
Interest	16	18	20	21	21	22
Total Outlays	210	231	244	258	281	311
Receipts	198	218	236	257	283	308
Surplus	—12	—14	—7	—1	2	—3
F.E. Rcpts.	212	230	243	258	282	311

tions for the two budget situations; tables 3B and 4B show the per capita real expenditures associated with each budget.

Table 5 shows the difference in per capita real (in constant 1969 dollars) expenditures that follow from the choice between the two budgets. Summing over

the five-year planning horizon, the cost of the \$225 increase in real per capita defense spending is approximately \$50 in housing, \$35 in education, \$30 in general government, \$35 in health and \$40 in transportation, \$15 in the consumption of basic necessities (food, clothing and housing),

TABLE 3—ALLOCATION RESULTS WITH HIGH DEFENSE BUDGET

Table 3A—Allocation of GNP by NIA Function and Year (%)						
Function	1970	1971	1972	1973	1974	1975
Housing	3.7	4.6	4.6	4.6	4.9	5.1
Education	6.4	6.5	6.6	6.6	6.7	6.8
Gen. Govt.	3.2	3.3	3.4	3.4	3.5	3.6
Health	6.6	6.8	7.0	7.2	7.4	7.6
Transportation	10.1	10.2	10.3	10.3	10.4	10.5
Defense	7.7	7.2	7.2	6.9	6.6	6.2
Investment	11.8	11.8	12.1	12.2	12.2	12.2
Other	8.9	8.5	8.2	8.2	8.1	8.0
Food, clothing, rent	41.7	41.1	40.7	40.5	40.2	40.0

Table 3B—Real per Capita GNP by Function and Year (1969 \$/Person)						
Function	1970	1971	1972	1973	1974	1975
Housing	149	192	199	209	229	249
Education	273	293	308	323	340	359
General Government	127	138	146	156	166	177
Health	277	300	321	343	367	393
Transportation	468	494	518	544	570	599
Defense	328	320	335	336	332	327
Investment	531	557	595	621	647	678
Other	406	400	405	421	431	445
Food, clothing, rent	1961	2012	2079	2155	2225	2308
Total	4521	4705	4907	5108	5306	5536

and \$15 in other goods. The target population is the more relevant denominator for each function. Thus, the \$35 per capita increase in resources available to education over the five-year period could translate into a \$350 expenditure for each child who enters first grade during this period.

Also, the higher budget causes additional inflation (0.3% over the five-year period). These are the true "costs" of the additional defense expenditures that should be weighed against the "benefit" of additional national security and the lower unemploy-

ment rates in 1971 and 72 (0.2% and 0.1%).

The numbers used in this analysis are admittedly crude. Yet, they do provide one measure of the opportunity costs involved in a major budget decision and thereby provide some basis for evaluating its combined allocation and stabilization effects. The numbers indicate that fiscal and monetary policy should not be determined only on the basis of short-run stabilization objectives but should consider the longer-run resource allocation that will result.

TABLE 4—ALLOCATION RESULTS WITH LOW DEFENSE BUDGET

Table 4A—Allocation of GNP by NIA Function and Year (%)						
Function	1970	1971	1972	1973	1974	1975
Housing	3.7	4.7	5.2	5.0	5.0	5.0
Education	6.4	6.5	6.7	6.7	6.9	7.2
General Government	3.2	3.3	3.4	3.5	3.6	4.0
Health	6.6	6.8	7.1	7.2	7.5	8.1
Transportation	10.1	10.2	10.3	10.4	10.5	11.0
Defense	7.7	6.7	6.3	5.9	5.5	5.1
Investment	11.8	11.8	12.0	12.2	12.2	12.3
Other	8.9	8.6	8.3	8.4	8.3	7.6
Food, clothing, rent	41.7	41.2	40.8	40.7	40.4	39.6

Table 4B—Real per Capita GNP by Function and Year (1969 \$/Person)						
Function	1970	1971	1972	1973	1974	1975
Housing	149	196	225	225	236	246
Education	273	291	311	325	348	381
General Government	127	137	148	157	172	197
Health	277	298	322	344	374	419
Transportation	468	491	520	545	579	628
Defense	328	299	292	288	279	269
Investment	531	553	590	624	651	682
Other	406	404	409	433	445	427
Food, clothing, rent	1961	2005	2081	2168	2248	2293
Total	4521	4672	4898	5109	5332	5543

TABLE 5—CHANGE IN PER CAPITA REAL EXPENDITURES BY FUNCTION AS A RESULT OF USING THE HIGHER DEFENSE BUDGET (1969 \$/PERSON)

	1971	1972	1973	1974	1975	Total
Housing	-4	-26	-16	-7	3	-50
Education	2	-3	-2	-8	-22	-33
General Government	1	-2	-1	-6	-20	-28
Health	2	-1	-1	-7	-26	-33
Transportation	3	-2	-1	-9	-29	-38
Investment	4	5	-3	-4	-4	-2
Other	-4	-4	-12	-14	18	-16
Food, clothing, rent	7	-2	-13	-23	15	-16
Defense	21	43	48	53	58	223
Total	32	8	-1	-25	-7	7

DISCUSSION

CHARLES E. METCALF: Since I did not receive either paper until the night before this morning's session, I shall keep my comments brief. The Packer paper is asking appropriate questions when it recognizes the endogenous nature of many budget outlays and emphasizes the allocative effects of macroeconomic budget decisions. On the other hand, it is an arbitrary exercise of sorts, and there is no indication how reliable the results are compared to what would be generated by various existing econometric models. The experiment was essentially to compare two defense budgets, with nondefense expenditures being forced to adjust to maintain a balanced full employment budget. Since the distribution of adjustments within the government sector was arbitrary, little insight was gained concerning these allocative effects. The primary value of the Packer effort lies not in his specific allocation results, but as an illustration of the need to consider the effects on resource allocation of fiscal and monetary policy.

The paper by Budd and Seiders takes such a partial and exogenous view of inflation that I don't know how to react to their conclusion that the redistributive effects of inflation are "small." It appears that they achieve their result by assuming away many of the reasons for expecting a change in the income distribution.

My first concern is with the notion of inflation as something exogenous which can be examined in the absence of other influences. The relationship between inflation and the price adjustment coefficients on sources of income and net worth may well depend upon the manner in which the inflation is generated. Second, the adjustment coefficients are assumed to apply uniformly across the distribution. For instance, inflation is assumed to affect the wage income of all individuals proportionally. There is substantial evidence to suggest that neither wage rates nor employment rates adjust uniformly. Related to this is the author's assumption that inflation is assumed to be neutral with respect to its effect on unemployment. Their argument that the unemployment effect may be either negative or positive de-

pending upon whether the inflation is of the excess-demand or the "cost-push" variety also serves as a reason *not* to consider inflation independently of its origins.

The authors provide a highly useful piece of partial analysis by providing insight into the consequences of differences in the composition of income or asset types at different levels of income or net worth. Yet there is no guarantee that the *ceteris paribus* assumptions implicit in applying adjustment coefficients to a fixed composition of income are in any way consistent with what is held fixed in the estimation of the coefficients themselves. The partial nature of the analysis prevents me from accepting the conclusion that "the redistributive effects of moderate changes in the rate of inflation on income and wealth, although present, are relatively small."

In summary, we need more insight into the processes by which inflation is generated, and its consequences in light of these processes. Budd and Seiders attempt to analyze the partial effects of inflation without regard to its origins, but require the use of adjustment coefficients which are not independent of these origins.

JOHN L. PALMER: My remarks will be limited to the Budd-Seiders paper. As Drs. Seiders and Budd have noted, much has been said, but little rigorous analysis done, concerning the redistributive impact of inflation on household income and wealth. Its presumed arbitrary nature does make it an important policy issue which deserves greater attention. If inflation and unemployment can be traded-off, then the selection of an appropriate macroeconomic policy requires a thorough knowledge of the implications for both the efficiency of allocation and the distribution of income and net worth of the various possible rates of inflation and unemployment. A study such as this is a step in accumulating this knowledge.

Viewing the paper in this context leads one to emphasize what I believe to be its major limitation—its partial analysis approach. The authors acknowledge this, but do not discuss its important implications. They examine the

short-run redistributive impact of an unanticipated increase in the rate of inflation from a previously stable rate, holding constant the rate (level?) of unemployment. There is a controversy concerning the role of expectations and adjustment lags as they relate to the verticality and stability of the Phillips curve relationship; however, I believe there would be nearly unanimous agreement among economists that an increase in the rate of inflation that is unanticipated by the various agents participating in the labor market is likely to be accompanied by a change in the level and rate of unemployment. In this event the average real wage rate may remain constant, but the adjustment coefficient for wage and salary income will surely be different from one for many income and wealth classes since there will be changes in hours worked; and there are compelling reasons to lead me to believe that this effect will have a significant, if not major, impact on distribution which will swamp the "pure" inflation redistributive effects as viewed by broad income and wealth classes.

Indeed, I wish to take this argument even one step further and suggest that a more appropriate model is one in which inflation and unemployment are endogenous, jointly determined variables within a complete structural system. Here we would include among the exogenous variables those instrumental policy variables over which we have some direct control (such as monetary and fiscal policy) in order to analyze their influence on inflation and unemployment. Then, given our estimates of how a particular macroeconomic policy will affect jointly the rates of inflation (anticipated and unanticipated) and unemployment, analysis such as in this paper will be valuable in helping to predict the total impact on income and net worth distribution. Such a model should enable us to confront the fact that different types of inflationary processes, e.g., one initiated by a change in monetary as opposed to fiscal policy, may have different implications for the values of the various adjustment coefficients of some income and asset types even though the same unanticipated rate of inflation is generated by each process. The adjustment coefficients in this

study are estimated, in many instances, by averaging over the various kinds of inflationary experiences that have occurred in this country since World War II.

Now allow me to discuss the paper more specifically—staying within the limits the authors have set for themselves—and in particular the results displayed in Tables 3–7. Their interpretation would be greatly facilitated if information on net worth by income class, the relative importance of various types of income by income class, and the size of asset holdings by net worth were included. Also, if I recall correctly, the sample size of the SFCC is only approximately 2,500 households with a disproportionately greater representation of the upper income and wealth groups. If the standard deviations associated with the estimated income and wealth changes were reported, particularly for low income quantiles, we would know how much confidence to place in these data.

There is a particular problem associated with the pension, social security and other transfers types of income. The SFCC lumps all these income sources into one of two categories so the authors had to use other estimates to determine the relative contribution of each of the numerous specific types of income within the two broader categories. They applied these same weights across all income groups rather than differentiating by income levels, a procedure which they justify on the basis that these income types are concentrated among lower income groups. This seems to me too gross a simplification, especially since one would expect that social security payments and a significant number of pensions would be distributed among middle and high income groups. Judicious use of the Survey of Economic Opportunity data, which is a sample of over 30,000 households heavily weighted with low income and non-white groups, could alleviate the above problem while permitting a breakdown of income and wealth redistributive effects by age, color, etc. with low standard errors.

The greatest strength of the paper lies in the methodology, both theoretical and empirical, that the authors develop in estimating some of their adjustment coefficients. The

methodological appendix should prove a valuable source for anyone wishing to further pursue the important type of empirical work they have done. Of particular note is their treatment of interest rate adjustment to inflation. They have also convincingly added to the evidence arguing against an inflation-induced fall of real wages, in part by testing a more reasonable interpretation of the wage-lag hypothesis than has previously been done.

Finally, I would like to stress the importance for future analysis of redistributive effects of macroeconomic policy of a merger of the income and wealth concepts which is

briefly touched upon in this work. Generally, except where a fluctuation of transitory income causes an undue hardship to some types of households, our concerns about redistribution are (or should be) with reference to some permanent income concept. Such a concept is particularly necessary if the redistributive consequences of labor market effects are going to be integrated with "pure" inflationary effects in a more complete model. I would hope that future analysis on the redistribution question will move in this direction, incorporating some of the insights of this paper into it.

THE POLITICAL ECONOMY OF ENVIRONMENTAL QUALITY

Environmental Pollution: Economics and Policy

By ALLEN V. KNEESE*

*Director, Quality of the Environment Program, Resources for the Future, Inc.,
Washington, D.C.*

Any discussion of environmental pollution by an economist naturally starts with the concept of externalities which has occupied the attention of some members of the profession for a long time. But very recent theoretical and empirical work has developed an approach which, while growing out of and incorporating the externalities concept, is sufficiently distinct in emphasis and orientation as to be legitimately deemed new. I will call this new approach, which I think to be much more commensurate with the character and scale of contemporary environmental pollution problems, *the management of common property resources* approach.¹ This paper is devoted to a general exposition of it and its policy implications.

A characteristic of most of the postwar discussion of pollution-type externalities, which at the same time is both a strength

and a weakness, is its focus on two-party situations.² The strength of this discussion has been its rather deep examination of the specific allocative effects of externalities under various circumstances. For example, the concept of reciprocal externality has been examined [5], the implications of bargaining between the parties for the traditional tax solution have been explored [3], and the relation of property rights to the externalities phenomena and its allocative implications has been clarified [4].

The weakness of this discussion has been that it has dealt with an artificial case which, unless properly interpreted, lends itself to some highly misleading inferences concerning policy. The case dealt with is about as follows: two parties, equal in economic power and in possession of full information concerning their own and their adversaries' positions, confront an externality situation in an economy in which resources allocation is in every other aspect optimal. Moreover, the environmental medium through which the external effect is transmitted (air, water, the electromagnetic spectrum, an ecological system, etc.) is fixed in supply, or, if it

*I wish to acknowledge helpful comments on an earlier draft by William Baumol, Peter Bohm, Blair Bower, James Buchanan, Ralph d'Arge, Mason Gaffney, Edwin Haefele, Robert Haveman, Orris Herfindahl, John Krutilla, Mancur Olson, Ezra Mishan, and Walter Spofford. The acknowledgment of help should not be taken to imply agreement with my conclusions.

¹This is not to say that common property-type problems have gone unnoticed in the literature, particularly in connection with "common pool" problems associated with petroleum and water resources and with fisheries. This literature is referred to later.

²Two excellent review articles have recently appeared to which the reader can be referred for a quick overview of the most salient writings. These are by Mishan [12] and Turvey [17].

can be altered, this can be done only by the two participants. Inferences which might be, and sometimes have been, drawn from consideration of such a situation are that bargaining between private parties eliminates "Pareto relevant" externalities; that the imposition of a tax based on marginal external effects distorts the optimal solution which would otherwise be reached by bargaining; that mergers can be an effective general solution to externality problems; and, much more generally, that the force of external effects is always met in some manner by the counterforce of private economic interest. The inference which should be drawn is that the essence of externality problems is that the conditions of this simple two-party case are not met in reality.

As regards environmental pollution, the case under examination is indeed highly unrealistic in several particulars. First, parties involved in an environmental pollution situation are usually anything but "separate but equal" insofar as organization, power, and information are concerned. The typical situation is one in which one or more sources of pollution, usually associated with a well-organized economic interest, affect a large and diffuse group of parties whose individual interests are hit relatively little. For example, a major source of pollution may affect many hundreds of fishermen, no one of whom finds it worth his while to bargain or even generate information. The fishermen also find the costs of organizing as a group prohibitive.³ In such a context, bargaining does not occur. What is more, no signal that important values are being destroyed is produced. Furthermore, what meaning could one possibly attach to a merger solution in such circumstances?

³An excellent and innovative discussion of the economics of cooperation and the conditions under which groups will be formed and function is found in Olson [13].

An additional complication with important implications for policy is that it is often possible to enhance an environmental resource that is being degraded through the application of large-scale, collective measures. For example, low river flows can be raised by regulating reservoirs, the reoxygenation capabilities of watercourses of all kinds can be improved, and ecological systems can sometimes be altered to counteract degrading effects (as in the case of the introduction of the Coho salmon into Lake Michigan). Usually, individual self-interest cannot be relied on to undertake, to an optimal degree, such activities which are characterized by the public good aspects of jointness in supply and/or exclusion problems. Finally, and most unfortunately from the point of view of the complexity of the problems, environmental pollution is not an isolated or random phenomenon, but in developed economies, at least, it must be regarded as inherent in man's production and consumption activities.⁴ Thus, an analytical orientation which regards externalities as a systematic and pervasive phenomenon in modern economies is much more on all fours with the real situation than one which regards them as a somewhat freakish anomaly which can often, if not usually, be efficiently controlled by ad hoc measures or private bargaining.⁵

In this connection it is useful to call attention to a simple fact, but one that seems to have been neglected by economists when reasoning about allocation problems. When materials—minerals, fuels, gases, and organic materials—are

⁴For development of this point, see Kneese, Ayres, and d'Arge [9].

⁵There is even something whimsical about the examples which are used in discussing the two-party cases. Bees and apple orchards have been with us a long time. A more recent favorite is sparks from a railway engine igniting a farmer's field—this being some eighty years after the introduction of the spark arrester and twenty years after the abandonment of the steam locomotive.

extracted and harvested from nature and used by producers and consumers, their mass is essentially unaltered in these processes. Material residuals are generated in production and consumption activities, and their mass must be about equal to that initially extracted from nature. Accordingly, it is basically depletive to speak of the consumption of goods. The services which material objects can yield are used, but their substance remains intact.

An important implication flows out of this for the allocation of resources within our institutional setting. While most extractive, harvesting, processing, and distributional activities can be conducted relatively efficiently through the medium of exchange of private ownership rights, the return of residual mass to the environment is heavily to *common property resources*, like the air and water, where the processes of private exchange cannot be expected to assign accurate relative values to alternative uses of the resource. In fact, we know from a relatively well developed body of theory concerning special types of common property resources like fisheries that perverse incentives lead to their overuse and misuse in both a static and dynamic sense.⁶ Thus, now that many properties of the natural environment have clearly become scarce resources, we are confronted with a vast asymmetry in the ability of our property institutions to form the basis for efficient resources allocation. What's worse, unless decisive countermeasures are taken, leading to more technically efficient use of materials and energy, we can expect this problem to grow rapidly more intense. If our spotty knowledge of past history is any guide, and residuals discharge is uncontrolled, we can expect residuals discharge to the environment to grow faster than national output

and the occurrence of public exposure to high residuals *concentrations* in the environment to grow faster still.⁷

What significant general conclusions can be drawn from the above discussion? I feel the following are important:

1. The discussion of two-party externalities has led to some important clarifications of the concept of externality, but because of the unreality of its assumptions its direct utility for policy prescription about pollution situations is small.

2. There is a need to reconsider general allocation theory with a view to the pervasiveness of externalities and to similarly rethink or adapt growth theory. Along with this, a deep reexamination of our national income accounting system is also called for. The profession has now begun to do these things.

3. We are particularly in need of operational analytical methods (and associated data) that can take account of broader systems of externality-type interdependencies than those that have typically been discussed in the economic literature.

The remainder of this paper is devoted to a discussion of some of the most pertinent efforts which have been undertaken under the last point and what I take to be the major policy implications flowing from them. The general orientation of these efforts is to view the various aspects of the natural environment as multiple purpose-multiple user natural assets, owned in common, which must be managed through some collective choice mechanism if they are to be developed, used, and conserved efficiently.⁸

⁷ The reason for the latter is explained in Ayres and McKenna [1]. In general, it is that exposure tends to be log-normally distributed, and with this distribution extreme concentrations will grow faster than mean concentrations as discharge increases.

⁸ Incidentally, it may be noted at this point that what might be called the "reciprocity of externality" argument presented by Coase [4] and others appears in a somewhat different light in the context

⁶ An early and excellent treatment is by Gordon [7]. A large literature was spawned by this article.

The kind of rethinking about pollution-type externalities which has taken place recently emphasizes that the generation of residuals is inherent in the production and consumption processes, that the types and amounts of residuals generated are strongly influenced by the production and consumption technologies adopted, and that residuals treatment (such as waste water treatment) does not reduce the mass of residuals but only changes their form. One of the strong lessons taught by this approach is that in a situation where the common property resources of environment are basically unpriced, externalities tend to be pervasive in the economy so that either general equilibrium or "second best" approaches are indicated. Another lesson is that the quantities of residuals discharged to air, water, or land are highly interdependent so that treating one environmental medium, such as water, in isolation from others can lead to undesirable secondary effects. (For example, the treatment of waste water leads to solid residuals which often are incinerated or otherwise pollute another environmental medium.) Still, our ability to produce operational general-equilibrium approaches to take account of these factors is as yet highly circumscribed, and we can gain important insights about intelligent policy approaches from more partial analysis.

In the next section of the paper, I will report briefly on several case studies of water quality which take the management of common property resources approach. They illustrate several points of importance to the later discussion. (1) Methodological tools exist which are useful for the quantitative study of environments as multiple purpose-multiple user natural assets. (2) These tools can be used to study the tax or charges approaches usu-

ally favored by economists as well as the more usual direct regulation approaches. (3) The enhancement of environmental common property resources has proved desirable on efficiency grounds in real cases. More generally, the ability to analyze and implement a wide range of technical measures has been shown to substantially improve efficiency. (4) Our present collective action institutions are ill-suited to realize these gains.

Following the discussion of these cases, I will sketch a new operational residuals management model devised by my associates Russell and Spofford [15]. The availability of this model opens opportunities for much more comprehensive analysis of the efficiency and distributional aspects of residuals management than has heretofore been possible.

The final section of the paper discusses what I take to be the most significant implications for policy flowing out of the studies discussed, as well as of some others which can only be cited because of space limitations. Also because of limited space, I have chosen to stress the empirical and operational aspects of the approach rather than the theoretical underpinnings. Even so, I can only give some interpretations and must cite other references for fuller discussions. I can only ask understanding for my oversimplification from those who understand the details of the studies which I sketch.

I. *Some Cases*

The first applied economic study to view an environmental problem as a matter of managing a multiple user-multiple product natural asset was a water quality study of the Delaware Estuary. This study, which had some substantial deficiencies, though it pioneered in important ways, was conducted by the Federal Water Quality Administration within a framework developed by economists. The

of the common property asset utilization than in the two-party direct interaction approach to externality. Cf. Baumol [2].

deficiencies were mainly in the narrow range of technical alternatives considered for improving water quality, and its neglect of the interdependency of various environmental media. Its achievements were to innovate some important methodological approaches and to do a competent, if limited, benefit-cost analysis of alternate management strategies including effluent charges or taxes.

Let me begin by sketching one aspect of the methodological approach which was particularly ingenious and which we will need to refer to again further on. The model of transfers of materials from dischargers to receptors introduced in the next section is one of the central elements in the realistic analysis of multiple-party pollution cases (see, for example, the discussion of the Russell-Spofford model below).

The Delaware Model

The Delaware Estuary was divided into thirty reaches, and the so-called Streeter-Phelps oxygen balance equations were adapted and applied to these interconnected segments. This led to a system of linear first-order differential equations. The transfer functions, which relate the change in concentration of a residual in segment i to an amount of input of the residual in segment j ,⁹ fortunately simplify to a set of linear relationships if steady-state conditions are assumed, i.e., if it is assumed that the waste discharge rate once established is constant and temperature and river flow are taken as parameters. In fact, the transfer functions can be represented by a matrix of coefficients.

⁹The transfer relations are fairly complicated because the degradation of organic wastes in the watercourse consumes D.O., whereas aeration, or movement of oxygen across the interface between air and water, tends to counteract this effect. Important variables that affect the oxygen balance for a given waste discharge are temperature and various characteristics of the watercourse.

The general name for such mathematical models of the movement and transformation of materials or energy in the environment is "diffusion models." This result is very useful because it means that these coefficients can readily be incorporated into a set of linear constraints which fit the linear programming format quite straightforwardly.¹⁰

A linear programming model similar to that sketched in the previous footnote was used to find the minimum costs for the whole estuary of different targets for water quality improvement given different management approaches, for example, ac-

¹⁰ Assume that the watercourse consists of m homogeneous segments and c_i represents the improvement in water quality required to meet the D.O. target in segment i . The target vector c of m elements can be obtained by changes of inputs to the water resource from combinations of the n segments. Define a program vector $x = (x_1, x_2, \dots, x_n)$ in which the values of x refer to the weight of waste discharges in each of the estuary reaches. In a feasible solution, these values represent the waste discharges at the various points which meet the target vector c . This vector generates D.O. changes through the mechanism of the constant coefficients of the linear system already described— a_{ij} = D.O. improvement in segment i per unit of x_j , $i = 1, \dots, m$; $j = 1, \dots, n$; and, of course, $x_j > 0$. If we let A be the $(m \times n)$ matrix of coefficients, then Ax is the vector of D.O. changes corresponding to x .

Now, recalling that c is the vector of target improvements, we have two restrictions on x , namely, $Ax \geq c$ and $x \geq 0$. The reader will have noticed that mathematically these are sets of linear constraints such as those found in a standard linear program. All we need is an objective function to complete the problem. Let d be a row vector where d_j = unit cost of x_j , $j = 1, \dots, n$. Notice that this assumes linear cost functions (although in more sophisticated models this is not necessary). We can now write the problem as a standard linear program,

$$\begin{aligned} \min \quad & dx \\ \text{s.t.} \quad & Ax \geq c \\ & x \geq 0 \end{aligned}$$

Of course, the transfer coefficients relate to a steady-state condition and to specified conditions of stream flow and temperature. Thus, the model is totally deterministic, and the variability of environmental conditions is handled in this analysis by assuming extreme conditions usually associated with substantial declines in water quality. The actual programs needed to solve the problem encountered in the Delaware Estuary were somewhat more complicated. The reader interested in the details should consult the paper by Sobel [16].

tual cost minimization for the whole estuary, cost minimization by zones in the estuary, uniform cutback in discharge at all outfalls, and effluent charges. The uniform cutback procedure might be regarded as a conventional administrative approach. The study found that for the water quality target, which benefit-cost analysis and an associated process of political decision-making indicated to be optimal, there was a very large cost difference (almost a factor of two) between this conventional approach of uniform cutback and cost minimization. Further, it was found that a single effluent charge of about 10¢ a pound for BOD would achieve the target level of water quality improvement at a real resources cost only slightly higher than the programmed cost minimization procedure. If the minimum cost solution were to be achieved through direct regulation, detailed control of each outfall would be required, and this would give rise to an extremely unequal cost distribution as among dischargers. The result that the effluent charge can approach a minimum cost solution is not very surprising to an economist because the charge tends to equalize marginal costs at different outfalls.¹¹

¹¹ This is strictly a necessary condition for cost minimization only when a homogenous "lump" of assimilative capacity is being allocated—or more formally, when all the coefficients in the transfer matrix are identical. When they are not, cost minimization requires that prices be "tailored" for each outfall. This explains why the solution based on a single charge only approaches but does not reach the programmed cost minimization solution. How closely it will approach is an empirical question relating to the magnitude of the a_{ij} 's. To see this, assume two dischargers with the following cost functions for reducing waste discharge:

- (1) $c_1 = f(x_1)$
 (2) $c_2 = f(x_2)$

where

x_1 = waste discharged from plant #1.

x_2 = waste discharged from plant #2.

c_1 and c_2 = cost of reducing waste discharge.
 Assume reach #6 is the "critical" reach. Therefore,

Later studies, using similar methodological tools and developing others, have significantly extended the results of the Delaware study in terms of water quality as a common property resource management problem. In general, the most significant result from these studies has been to demonstrate beyond a scintilla of a doubt that technical measures to improve the quality of the water resource, such as flow regulation via reservoir regulation, or groundwater pumping, mechanical reaeration of streams, and other regional-scale collective measures, can greatly enhance

- (3) $R_6 = a_{61}x_1 + a_{62}x_2$ (i.e., "binding" constraint)
 R_6 = the "standard"

Form the Lagrangian,

- (4) $L = c_1 + c_2 + \lambda(R_6 - a_{61}x_1 - a_{62}x_2)$

At optimum,

- (5) $\frac{\partial L}{\partial x_1} = \frac{dc_1}{dx_1} + \lambda(-a_{61}) = 0$

- (6) $\frac{\partial L}{\partial x_2} = \frac{dc_2}{dx_2} + \lambda(-a_{62}) = 0$

If we include the constraint $R_6 = a_{61}x_1 + a_{62}x_2$, we have 3 equations and 3 unknowns (x_1 , x_2 , and λ).

Solving for λ , we get:

- (7) $\lambda = \frac{1}{a_{61}} \frac{dc_1}{dx_1}$

and

- (8) $\lambda = \frac{1}{a_{62}} \frac{dc_2}{dx_2}$

Note that:

$$\frac{1}{a_{61}} \frac{dc_1}{dx_1} = \frac{1}{a_{62}} \frac{dc_2}{dx_2}$$

or

$$\frac{dc_1}{dx_1} = \frac{a_{61}}{a_{62}} \frac{dc_2}{dx_2}$$

Note also that $\lambda \neq 0$ unless either dc_1/dx_1 or $dc_2/dx_2 = 0$. Because both equations (5) and (6) are equal to zero, they may be set equal to each other:

$$\frac{dc_1}{dx_1} = \frac{dc_2}{dx_2} + \lambda(a_{61} - a_{62})$$

(I am indebted to my associate Walter Spofford for developing this demonstration.)

the efficiency of a water quality management program beyond what would be possible if waste water treatment were the only alternative included.¹²

Another generalization supported by recent economic-engineering research, mostly still unpublished, is that even at points of waste generation many technical alternatives besides waste treatment can be efficiently used to reduce waste discharges. Industry studies and evidence from those instances where sewer surcharges have been levied on industries indicate that process redesign, recycling, and by-product recovery are fully as important in reducing industrial residuals discharges as treatment in a conventional sense and frequently much less costly. These results apply not only to waterborne but to solid and airborne residuals as well and are important for our discussion of the Russell-Spofford model in the next section and the concluding considerations of policy implications.

II. *The Russell-Spofford Model*

We have learned much from studies viewing water pollution as a common property resource management problem. The results are beginning to have an important effect on policy formation at several levels of government, both in the United States and abroad, as well as on education in the pertinent parts of engineering and economics. But as is implied by the introduction to this paper, they still go only part of the way toward analyzing the common property resources of the environment as multiple purpose-multiple

user natural assets with technical and economic interdependencies often existing among different common property resources.

As was pointed out in the introduction, recent literature has yielded the insight that residuals management confronts economics with a new type of general-equilibrium problem. Also, it was indicated that some pertinent literature reflecting this has begun to appear. Still, we are far from being able to do a satisfactory general-equilibrium analysis of the problem. But we are moving rapidly toward implementation of what might be called a partial-general model which retains many of the most important insights from the general-equilibrium approach. An operational model has been devised by my associates Russell and Spofford [15] which we will soon apply to an actual region—probably the Delaware estuary area.¹³ I discuss here only a highly skeletonized version of the model.¹⁴ Its essential components are as follows:

A linear programming inter-industry model that relates inputs and outputs of the various production processes and consumption activities at specified locations within a region, including the unit amounts and types of residuals generated by the production of each product; the costs of transforming these residuals from one form to another (as gaseous to liquid in the scrubbing of stack gases); the costs of transporting the residuals from one place to another; and the cost of any final discharge-related activity such as landfill operations.

The interindustry model permits choices among production processes, raw material input mixes, by-product produc-

¹² The most important published study of this type is by Davis [6]. A short interpretation of this study along with several other pertinent ones is contained in Kneese and Bower [10]. Other, unpublished studies supporting the above conclusions are of the Miami Basin in Ohio, the Raritan Bay, and the Wisconsin River Basin. All of these have been cooperative studies by economists and engineers.

¹³ A rather full discussion of the model is found in Russell and Spofford [15] in Kneese and Bower [11].

¹⁴ This exposition is based on unpublished material prepared by Clifford S. Russell and Walter O. Spofford.

tion, recycle of residuals, and in-plant adjustments and improvement, all of which can reduce the total quantity of residuals generated. That is, the residuals generated are *not* assumed fixed either in form or in quantity. This model also allows for choices among transformation processes and hence among the possible forms of the residual to be disposed of in the natural environment and, to a limited extent, among the locations at which discharge is accomplished.

Environmental diffusion models (atmospheric diffusion, stream diffusion, and noise diffusion models) are used to relate the amounts and types of residuals discharged into the natural environment—both to the atmosphere and to watercourses—to the amounts (i.e., ambient concentrations) and types of residuals that are present at the various receptors—man, animals, plants, and inanimate objects.

In the case of noninteracting residuals, the total quantity (i.e., concentration) of a given residual at any point in time and space is a linear sum of the contributions from all sources. This has already been explained in connection with the Delaware estuary model. Air and noise diffusion models work similarly. Thus we can express the ambient concentrations at various receptors r as the set of linear equations $Ax=r$ where A is an $m \times n$ matrix of transfer coefficients.

The quantities of residuals discharged to the environment, x_j , $j=1, \dots, n$, are given as an output of the interindustry LP model, and the transfer coefficients, a_{ij} , $i=1, \dots, m$; $j=1, \dots, n$, are computed from environmental diffusion models. The vector of ambient concentrations of the various residuals at all the receptor locations, i , $i=1, \dots, m$, due to residuals discharged from all sources, x_j , $j=1, \dots, n$, is the output of the above system of equations.

Receptor-damage functions use the vector r as input. Where damage functions are not available, ambient standards may be used. For this situation, the above equation set may be modified to include the standards

$$Ax \leq S$$

where

S_i , $i=1, \dots, m$; = ambient standards. Since this is a linear constraint set, it can be transferred in its entirety to the inter-industry LP model. The duals of the maximization solution will then indicate the implied marginal value of the constraints in terms of the objective function. When damage functions are available, it is possible to compute marginal damages at the locations of residual discharge for all sources.¹⁵ As described in the footnote, these prices reflect the sum of marginal costs imposed on all receptors by each discharge point.

¹⁵ It is assumed that the functions are continuous and have continuous first derivatives. For this case, the damages at each location, D_i , $i=1, \dots, m$, may be written as a function of the corresponding ambient residual concentrations, R_i , $i=1, \dots, m$, or

$$D_i = f(R_i) \quad i = 1, \dots, m.$$

For each residual, the total damages to all receptors and/or uses in the region is given by:

$$D_T = D_1 + D_2 + \dots + D_m = \sum_{i=1}^m D_i$$

Since the actual damage, D_i , is a function of the ambient concentration, R_i , the following expressions for total marginal damages obtain:

$$\begin{aligned} \frac{\partial D_T}{\partial x_1} &= \frac{dD_1}{dR_1} \cdot \frac{\partial R_1}{\partial x_1} + \dots + \frac{dD_m}{dR_m} \cdot \frac{\partial R_m}{\partial x_1} \\ &\vdots \\ \frac{\partial R_1}{\partial x_n} &= \frac{dD_1}{dR_1} \cdot \frac{\partial R_1}{\partial x_n} + \dots + \frac{dD_m}{dR_m} \cdot \frac{\partial R_m}{\partial x_n} \end{aligned}$$

However, $\partial R_1 / \partial x_1 = a_{11}$, $\partial R_2 / \partial x_1 = a_{21}$, $\partial R_m / \partial x_1 = a_{m1}$, and so on. Thus, total marginal damages resulting from each discharge are obtained by post multiplying the vector of derivatives of damages with respect to concentrations by the transpose of the A matrix, A^T . The marginal damages are evaluated using this equation set and the resulting vector $\partial D_T / \partial x_i$, $i=1, \dots, n$, is returned to the inter-industry LP model as prices on the residuals discharged to the environment.

When an iteration is completed using these prices, a new discharge vector results and the process is repeated with newly computed prices. Of course, each of the iterations involves a linear approximation to a response surface which is not linear. Thus a problem of determining step sizes permissible in each iteration arises. It is not appropriate to pursue this interesting problem here.

Enhancement of the common property resource through such measures as river flow regulation and mechanical reaeration is also included in the model in ways which cannot be discussed in this short paper. The next step is to try to take broader ecological considerations explicitly into account in the model.

In sum then, the Russell-Spofford model permits analysis of numerous strategies for the management of common property environmental resources in a region. These include charges and several types of standards to induce private and public investment in numerous types of residuals control technologies and techniques for the enhancement of common property resources. The analysis can consider all the pertinent interdependencies among liquid, gaseous, and solid residuals in a region.

III. Policy Discussion

General-Central Problems for Policy

My reading of the literature of the common property assets approach to environmental problems suggests two central problems for appropriate policy formation. The first is the development of systematic and routine procedures for the pricing of common property environmental resources. I think that economic research has already shown that acceptably rational pricing of these resources is by no means a hopeless task (see the Delaware study, for example). Of course, in many cases this may mean an infinite price, i.e.,

a complete prohibition of the use of the resource for a particular purpose. For example, I find it difficult to imagine that the social costs of forbidding the discharge of a heavy metal like mercury to watercourses could outweigh the social benefits. In those cases where there is great uncertainty concerning the effects (especially the longer-term ones) of using the environmental resource, a much greater burden of proof should be placed upon the proposed user than is now the case. In many instances at the present time, we do not even know what users are using our common property resources for. The apparent ignorance of the regulatory scientific community concerning the discharge of mercury to watercourses in the United States is a striking case in point. We are still operating largely under property laws which impose the burden of proof on the damaged party. This legal posture at least implicitly assumes that rapid growth is a central public objective and that external costs will be small—both are propositions which may be seriously questioned in our contemporary economy.

The second main problem for policy indicated by the common property assets approach is that of institution building. Environmental pollution problems are inherently regional. Unless they are global (such as CO_2), they occur in a watershed, an air-shed, or some other problem shed which almost never coincides with existing governments of general jurisdiction.¹⁶ It is within such problem sheds that external effects occur and that the opportunities for reducing residuals generation, transformation from one form to another, or al-

¹⁶ This is not to say that existing governments of general jurisdiction are never appropriate to the management task or could not take important initiatives of a "second best" character. Some problems are best handled at the national level, and the states have tremendous powers (in principle) which they mostly don't use.

tering the environmental assimilative capacity occur. In such regions there is a technical job of analyzing and implementing management strategies as well as a political job of making collective choices about a collective resource. Our present structure of federal, state, and local governments is quite unsuitable for implementing an efficient or even coherent approach for managing common property assets—or even for a single one like water.

But we are, as of this moment, in a better position in terms of research base to suggest a strategic approach to the water quality management problem than to the other environmental management areas. Also, there has developed, in bits and pieces, a national policy approach to this problem which is somewhat more definite than in the other areas. Thus, we can contrast the actual (implicit) strategy with the one which research such as that reported in earlier sections suggests would be preferable. Accordingly, I will emphasize the water quality problem rather than present a balanced discussion of the whole range of environmental pollution problems.

A Partial Policy Program—Water

To start with, I would like to characterize briefly the present strategy of the federal government for achieving water pollution control in the United States. This strategy is based on two main elements. The first is financial support for municipal waste treatment plant construction. This program has been going on for a long time, and presently it is possible for municipalities to cover up to fifty-five percent of the costs of waste treatment plant construction from federal grants. The second main element in our national strategy is enforcement actions against individual waste dischargers. I think it is fair to say that the results of this strategy have been

disappointing. Municipal treatment plant construction has been lagging partly because federal appropriations for treatment plant construction have fallen far behind authorization, and it is clear that municipalities are holding up construction until federal funds become available. Many plants that have been built are not operating well. The difficulty and cost of mounting effective enforcement proceedings, as well as the political power of the larger industries, has effectively stymied enforcement at both the federal and state levels. More generally, our record of trying to impose direct federal regulations on large industries has, of course, been dismal.

Under our present system, a certain amount of subsidy is also available to industrial plants. Some are connected to municipal systems and can benefit from the subsidies to municipal treatment plant construction. Furthermore, five-year tax amortization of pollution control facilities is available. This is a particularly perverse form of subsidy because it cannot help those marginal firms which often serve as the excuse for subsidy arrangements. Tax writeoffs have the effect of providing most assistance where it is not particularly needed and, unless counteracted by other provisions, letting the industrial plant where assistance might be justified die.

The result of this strategy has been a continuing deterioration of water quality due to a low rate of municipal treatment plant construction, poor design and operation of existing plants, and in many basins the overwhelming growth of industrial waste discharges.¹⁷ Progress toward institution building for effective and efficient water quality management has been nil.

An alternative strategy for dealing with our national water pollution problems grows rather naturally out of the eco-

¹⁷ In this connection, see the GAO Report [8].

conomic research which has been accomplished on the matter. This strategy is also based on two main elements. The first rests on the concept that the waste discharger should insofar as possible bear the costs his waste disposal activities impose on the common property assets of society, and the second recognizes that in many of our highly developed basins, where pollution problems are concentrated, great gains in effectiveness and efficiency can be obtained by the implementation of a systematic and well-integrated water quality management plan on a regional basis. The latter would contain elements other than just the treatment of waste waters at particular outfalls.

The present subsidy arrangements are quite different in their impacts than the effluent charges system. First, the system of effluent charges is based on the concept that payment for the use of valuable resources is necessary for efficiency, whether they happen to be privately or collectively owned. These payments will affect industrial producers' decisions to generate and discharge residuals. They will also be reflected in the price of intermediate and final goods so that a broader incentive will be provided to shift to goods with a lesser environmental cost. The present system of subsidies has the unfortunate effect of subsidizing those goods most, the production of which, directly or indirectly, makes the heaviest use of common property assets.

Second, to the extent that the subsidy-enforcement¹⁸ system is effective in reduc-

¹⁸ Enforcement must be linked to subsidies because even though the latter are often erroneously called incentives, they do not in themselves provide an incentive to take action to control waste discharges. Even if an industry is paid a major proportion of the cost of waste treatment plant construction, it is still cheaper, from the point of view of the industry, to dump untreated waste into the river. Thus, the subsidy arrangement cannot work unless accompanied by enforcement or other pressures on the waste discharger.

ing discharge, it tends to bias the choice of techniques in an inefficient direction since as a practical matter only treatment plants qualify for subsidies. This provides an incentive to construct treatment plants with federal subsidy, even though internal controls would be cheaper as in many, if not most, instances they are.

Finally, the effluent charges system yields revenue rather than further straining and eroding an already seriously over-extended tax system. This revenue can be put to useful public purposes including improvements in the quality of our environment.¹⁹

Industry has been emphasized but municipalities too are paying only part of the social costs associated with the wastes they generate, and what they pay is rather capriciously distributed depending on how much waste water treatment they have implemented and whether they have qualified for federal subsidies. The effluent charges system would give these municipalities an incentive to proceed expeditiously in the effective treatment of waste. Our present policies put heavy emphasis on the construction of plants with little or no follow-through on operations. The effluent charges system focuses on what is put in the stream and thereby offers an incentive for effective operations of existing facilities.

Despite what appear to be compelling reasons for favoring the effluent charges system as one of the cornerstones of effective and efficient regional water quality

¹⁹ From an economic point of view, perhaps the best imaginable tax base is an activity that causes external diseconomies. Not only does a tax on such a base yield revenue, but it tends to improve the overall allocation of resources, thus yielding an "excess benefit." It should be noted that the yield from an optimal effluent charge would only be equal to what should optimally be spent on common property resource enhancement only under special circumstances. Usually it would be more or less. For an analysis of this, see the Appendix to chapter 10 (by J. Hayden Boyd) in Kneese and Bower [10].

management, it may be difficult for particular states and regions to pioneer such a substantial departure from previous practice, although surprising initiatives have been taken in several states and regions recently. Nevertheless, the federal government's greater insulation from powerful local interests provides an opportunity for leadership. One approach would be for the federal government to levy a national effluent charge on all waste discharges above some minimum amount. This charge could be considered a minimum which could at their discretion be exceeded by a state or regional agency having responsibility for water quality management. Revenues obtained by the federal government could be made available for purposes of financing the federal program with the excess turned over to other governments of general jurisdiction or, and I think preferably, the revenues could be used to establish regional water quality management agencies which are the other element in the proposed strategy. The rationale for this element has been rather fully discussed earlier.²⁰

One way for the federal government to encourage regional agencies would be to establish incentives and guidelines for the organization and operation of regional management agencies, either under state law or through interstate arrangements. An agency with adequate authority to plan and implement a regional water quality management system would be eligible for a grant of funds to support a portion of its budget to help staff the agency and to make the first data collections, analyses, and formulation of specific measures for water quality management. If the federal government is satisfied that the proposed program and the plan for its implementation satisfy criteria for its efficient

operation, the agency might be eligible for a grant to assist it with actual construction and operating expenses. Such a system might appropriately be limited to the early implementation—say, five years. During this period, it would be necessary to work out longer-term arrangements for financing the agency. Clearly, the proposed effluent charges system could play a major role in this. Presumably, administration of the effluent charges system would be turned over to the regional agencies with the federal level of charges continuing to be regarded as a baseline. In this manner, regional scale measures for the management of the common property asset would be financed while at the same time providing incentives to waste dischargers to cut back on their emissions. Special provisions might be included in the federal law toward marginal industrial plants which might go under and where there is a broader social interest in protecting them. It should be noted that where serious efforts to implement regional water quality management have been undertaken (as in the Delaware and the Miami), one of the most serious problems has been to set up adequate financing arrangements. Once the regional agency is functioning, collection of the effluent charge would be turned over to it. It would be expected to improve and refine the charges system.

There is some encouraging evidence that a strategic reconsideration of our approach to water quality management is occurring. In both the executive and legislative branches of our national government, serious discussion of approaches emphasizing regional management and economic incentives is occurring. Similar developments are also happening at lower levels of government.

Interestingly enough, the main opposition to the approach outlined here—especially the effluent charges part of it—has

²⁰ The elements are embodied in a bill before Congress introduced by Senator Proxmire [4].

come from industry and the so-called conservationists. The industrial opposition stems, I believe, from the recognition that effluent charges would be effective and not subject to the delays and general manipulations which the administrative-enforcement approach has permitted. In the case of the conservationists, the opposition seems to reflect a puritanical ideal that common property resources should not be used for residuals disposal at all. This is idealism run wild, as a moment's reflection about the first law of thermodynamics will reveal. In their case, what they regard as the "best" is clearly an enemy of the good.

Other Steps

Even should we move in the direction of greatly improved water quality management, this is still a far cry from a complete and coherent approach to the management of our environmental common property resources. But there is at least some small ground for encouragement that economically rational strategic approach to environmental management may be developing. Under the Environmental Policy Act of 1969, the President appointed a Council on Environmental Quality. This Council was modelled on the Council of Economic Advisers. The first Council consists of three highly respected environmentalists who issued their first report in mid-1970. In it there is much evidence (far more than in any previous important public document) of recognition of the importance of the twin features of environmental quality management problems emphasized in the first part of this section—incentives and institutions. Also, the need for a coherent concept to guide environmental management is recognized. Many constructive recommendations for study and experimentation are made, but the report does not press for action to nearly the extent I believe would be appropriate.

Another development of 1970 of significance for improving our national organizational posture in regard to environmental problems is the creation of EPA—the Environmental Protection Agency. On the recommendation of the Commission on Reorganization of the Executive Branch of the Government (the Ash Commission), the President proposed, and Congress accepted, a reorganization of the federal pollution control agencies. In this connection, the Commission and other Executive Office advisory arms gave careful consideration to the rationale for a unified approach to environmental quality management described in earlier sections of this paper and which also underlies the Russell-Spofford model.²¹ EPA combines the pollution control and related research activities of the federal government pertaining to air, water, solids, pesticides, and radiation into a single agency reporting directly to the President. These functions were formerly performed in virtually complete isolation from one another and housed in diverse agencies such as HEW, Interior, Agriculture, and the Atomic Energy Commission. While anyone who has observed the Washington scene for a period of time has just cause to be skeptical of the efficacy of reorganizations, one can still recognize the convincing rationale of this one. It puts us in a better position to view residuals generation as an inherent part of our production and consumption activities rather than an isolated phenomenon and to recognize the interdependencies among types of residuals.

To sum up, the strategic policy approach developed with respect to the environmental resource where policy is most advanced is subject to fundamental criticism from the point of view of economic

²¹ Fortunately, there was ample opportunity to communicate these thoughts to the panel of the Ash Commission working on the environmental aspects of the reorganization.

theory and empirical research. On the other hand, evidence accumulates that approaches based on the latter are penetrating the thinking of governments at both the national and other levels. There are good grounds for hope that our strategy for water quality management can be somewhat reoriented and improved and that those policy areas pertaining to the management of common property assets which are still in their formative or undeveloped states can be based on sounder conceptual and policy research ground. This is the bright side. The other side is that progress is very slow, the problems are urgent, and we know that vested interests can often stave off really effective policy until they are ready to accommodate it.

REFERENCES

1. R. U. Ayres and R. P. McKenna, *Alternatives to the Internal Combustion Engine*. Baltimore 1971.
2. W. J. Baumol, "On Taxation and the Control of Externalities," unpublished paper.
3. J. W. Buchanan and W. C. Stubblebine, "Externality," *Economica*, Nov. 1962, 29, 371-84.
4. R. H. Coase, "The Problem of Social Cost," *Jour. of Law and Econ.*, Oct. 1960, 3, 1-44.
5. O. A. Davis and A. Whinston, "Externalities, Welfare, and the Theory of Games," *Jour. of Pol. Econ.*, June 1962, 70, 241-62.
6. R. K. Davis, *The Range of Choice in Water Management: A Study of Dissolved Oxygen in the Potomac Estuary*. Baltimore 1968.
7. H. S. Gordon, "The Economic Theory of a Common Property Resource: The Fishery," *Jour. Pol. Econ.*, April 1954, 57, 124-42.
8. U.S. General Accounting Office, *Examination into the Effectiveness of the Construction Grant Program for Abating, Controlling, and Preventing Water Pollution* (Federal Water Pollution Control Administration, Department of the Interior), No. B-166506, Nov. 1969, Report to the Congress by the Comptroller General of the United States.
9. A. V. Kneese, R. U. Ayres, and R. C. d'Arge, *Economics and the Environment: A Materials Balance Approach*. Baltimore 1970.
10. A. V. Kneese and B. T. Bower, *Managing Water Quality: Economics, Technology, Institutions*. Baltimore 1968.
11. ———, eds., *Environmental Quality Analysis: Research Studies in the Social Sciences*. Baltimore, forthcoming.
12. E. J. Mishan, "Reflections on Recent Developments in the Concept of External Effects," *Can. Jour. Econ. & Pol. Sc.*, Feb. 1965, 31, 1-34.
13. M. Olson, *The Logic of Collective Action*. Harvard 1965.
14. W. Proxmire, S. 3181, *The Regional Water Quality Act of 1970*, Nov. 1969.
15. C. S. Russell and W. O. Spofford, "A Quantitative Framework for Residuals Management Decisions," in Kneese and Bower [11].
16. M. J. Sobel, "Water Quality Improvement Programming Problems," *Water Resources Research*, Fourth Quarter 1965, 1, 477-87.
17. R. Turvey, "On Divergences between Social Cost and Private Cost," *Economica*, Aug. 1963, 30, 309-13.

DISCUSSION

KENNETH E. BOULDING: It is hard to take issue with a paper which hits so many nails so precisely on the head. I have participated for nearly four years now in an interdisciplinary seminar on environmental problems organized by Dr. Morris Garnsey at the University of Colorado, and the main thing I have learned in regard to environmental problems is the relative backwardness of the physical and biological sciences. It is frequently our ignorance of the physics and chemistry of the atmosphere, or of the biology of large ecosystems, which holds us up on the practical side of environmental improvement. The physical sciences are good at the very small and the very large, but when it comes to medium-sized systems like the earth they often seem strangely ignorant. Meteorologists do not really know whether man's activities are warming up the earth or cooling it down, and biologists do not really know the effect of DDT on the biosphere. It is in this atmosphere of scientific ignorance that the wilder speculations of the environmental eschatologists get taken as gospel, and many young people today seriously seem to believe that the earth will become uninhabitable in their lifetimes, never having heard of ultrastability in complex systems. This is not to deny, of course, the positive probability of irretrievable catastrophe, but I suspect it is much smaller than the doom-and-gloom school suggests.

By comparison with the ignorance and even obscurantism of the natural scientists, economics stands out like a clear beacon of eighteenth century enlightenment. It is true that we also have our lunatic fringe who virtually deny the existence of public goods and public bads and think that all things can be done by private bargains between smoky railroads and rational dairy farmers. Kneese has done a fine-tuned job of intellectual assassination on these types, such that no further comment is necessary. If one had to sum up what economists have to say on this matter in five or six propositions, the first would be that we produce bads (pollution) because bads are jointly produced with goods. The second is that negative commodities (bads) should have negative prices, and the third is that public goods and

bads require political organization if we are not to get too little of the first and too much of the second. A fourth proposition is that goods depreciate and bads appreciate, following the generalised and extended Second Law of Practically Everything, that all things slide down towards a middle muddle unless somebody does something about it, that is, unless information of the right kind is applied at the right time and place in the system. There is a fifth proposition that depreciation and appreciation are not continuous functions of use or load, but exhibit threshold or overload phenomena, which is what causes crises. Crises in the differential calculus are very rare. Unfortunately continuous functions, which are fine for celestial mechanics, are characteristic of social mechanics only over small ranges of variation, and most social problems arise because of discontinuous functions—the road that suddenly jams up as one more car appears on it, the river that refuses to clean itself up under a single addition of sewage, the international system that breaks down into war, or the city that erupts into riot, when one small straw is laid on some existing back. This is one reason why mathematical models of this problem should never assume continuous functions, but should always explicitly include some kind of step-function or limit of each variable at which the system breaks down. A sixth proposition, which I have been stating on and off for years without anybody paying much attention, is that human welfare is a state, condition, or stock rather than a flow variable, though it may be related to certain flow variables such as variety or quality of inputs such as food. What we think of as the environment, therefore, is simply part of the capital stock of human welfare and must be evaluated as such. Nature has no values, unless we count survival as a value; only man has the effrontery to try to calculate a net worth for the universe, or at least for that part of it which is relevant to him. Consequently, when we are evaluating the environment, we are also evaluating man himself as part of a total system. What we think of as "capital" must include not only human artifacts, and human bodies and minds, but also

the whole range of goods and bads with which man is surrounded, whether these are artifacts of man or whether they are "natural" objects. To this enormous list of goods and bads we must add a similar set of valuation coefficients or shadow prices in order to be able to add up the list into a single number. It is the setting of these valuation coefficients about which there is so much argument, especially in the case of bads which do not have a ready market. Certainly in calculating a net societal worth, or even a net national product, existing bads in the one case or produced bads in the other should be deducted from the total of values of goods. How to put a price on these bads, however, is the really ticklish question. What is pure air or water worth, to how many people? Where these are public goods it is very hard to estimate how much goods people are willing to sacrifice in order to get rid of how much in the way of bads.

One cannot help feeling that part of the confusion in this and other areas of social thinking arises from an unwillingness to recognize that negative commodities (bads) have a reality and properties of their own, and that they are not merely "minus goods." The simple algebra, indeed, in which minus minus is a plus, does not apply to social life. If a bandit abstains from killing me because I give him my money, this is not the same thing as an exchange of goods. The negation of a negation, Hegel and Marx notwithstanding, does not necessarily produce anything positive, but may simply multiply human misery. We need, indeed, an algebra for society in which $(-x)(-x) \neq +x$. Many of the problems of pollution arise because we do not seem to realize that doing away with a bad is not the same as failing to produce it in the first place—that cleaning up a river is not the same thing as not polluting it in the first place. Whether it pays, however, from the point of view of society, to abstain from producing a bad at the cost of not producing a good, or to produce "negative bads" (activity which might have increased goods being devoted to diminishing bads), or to produce additional goods in order to overcome the increase of bads or the diminution of goods, or to produce different goods which are not as good as those we might have produced but have fewer bads going along

with them as joint products—all this has to be evaluated through multiplying the quantities of each good or bad by the appropriate positive or negative shadow price and then summing the result.

The task of evaluating the universe (or even part of it) indeed is so formidable that we may need to reevaluate our evaluative strategy itself. If the best cannot be found, and if the demand for the best and nothing but the best turns out to be the enemy of the good, as Kneese suggests, what strategy of evaluation do we settle for? Do we "satisfice" and say anything above a certain level will do, no matter what? Do we simply aim to avoid disaster—to identify the cliffs and build fences around them? Or do we take the conservative strategy of trying to stay where we are, on the assumption that we are lost in a thick forest on a dark night, so that even if there were any better place than where we are now there would be no way of getting there without disaster?

These are tough enough problems even in the realm of private decisions and private goods. When we come to public goods and bads, the problems get even tougher. Public goods, we all agree, will not be provided adequately, and public bads will not be disposed of adequately, unless there is a "public," that is, unless the people who share the enjoyment of public goods and the distress of the public bads constitute enough of a community to permit public action. Just what constitutes a "public," however, and how it gets that way, is one of the most puzzling and at the same time the least asked question in social dynamics. It is astonishing how economists, for instance, especially public finance types, simply take the national state, or even its subordinate political entities, for granted, without even raising the question as to how things got that way or why things should stay that way. My own view is that the public goods problem is quite inseparable from that of the "grants economy"—that is, the structure and extent of one-way transfers; it is also closely related to systems of threat and systems of legitimation, both of which are barely visible in economic theory. The problem of the "public" is related to the creation of benevolence and malevolence—that is, interdependence of utili-

ties, which is why the concept of the Paretian optimum, which assumes selfishness, is almost useless except for passing examinations with. The real world is so shot through with a complex network of benevolence and malevolence that selfishness has to be seen as a most unlikely case—merely the zero on the malevolence-benevolence scale. It may be, indeed, that the externalities associated with the production of malevolence (a major industry of the human race) present a greater problem than all the domestic and industrial effluents of the world. This is why the international system, which is a spectacular producer of malevolence because it is operated mainly by organized threat, is probably the greatest polluter of all and the greatest threat to man and his environment. The reasons for this, however, are almost exactly the same as the reasons we give why the smaller problems of effluent are so difficult to solve in practice, in spite of the fact that the solutions are frequently known. The absence of a Delaware Valley “public” is the same sort of phenomenon on a smaller scale which the absence of a world public is on the international scale. Perhaps the last word to the environmental economist, therefore, is that the presence of pollution is a symptom of the absence of community.

JACK W. CARLSON: Allen Kneese’s paper contains the same advice he has consistently and patiently offered officials of the Federal Government for at least the last five years. The significant difference in this paper is his optimism that his advice is being taken seriously. He continues to recommend that the Federal Government make a greater commitment to manage air and water common property resources. He believes this is occurring, and I agree.

As evidence, Mr. Kneese refers to the creation of the new policy-oriented Council on Environmental Quality and the new management-oriented Environmental Protection Agency, and thereby the enhanced possibilities of residuals waste management. Based on the Federal Government’s actions last week, additional and more dramatic evidence is available. The Congress passed the Clean Air Act Amendments of 1970, which (a) require the Federal Government to establish national

ambient air quality standards during 1971, (b) require enforcement of the standards within three years thereafter by the states with a supporting role provided by the Federal Government, (c) establish very stringent federal emission standards for automobiles to be achieved by 1976, at the latest, and also emission standards for aircraft; (d) require regulation of hazardous materials in general and especially as inputs into fuels; and (e) provide for a one-year study of the actual and potential problems with noise and the need for regulation by the newly-created Office of Noise Control in the Environmental Protection Agency.

The President announced plans to broaden the provisions of the Refuse Act of 1899 from its sole concern with navigation to include water quality management. Industry must obtain a license to discharge into navigable streams and their tributaries. This applies to all new industrial discharge within six months and to existing industrial discharge within three years. States must certify that the anticipated industrial discharge is compatible with state water quality standards. The new Federal Environmental Protection Agency must establish guidelines for determining abatement of pollutants that it deems harmful, especially those pollutants affecting other states. The Corps of Engineers is required to manage the issuance of licenses and to supervise conformance with the provision of each license.

As the result of these initiatives and past public action, public policy affecting air and water quality management is revealed:

(1) There is a major public commitment to manage the common property resources of air and water quality.

(2) The governmental entity that has been chosen to take the policy *lead* is the Federal Government.

(3) The Federal Government has taken the *management* responsibility for pollutants from *mobile* sources (automobiles, aircraft, etc.).

(4) The governmental entity that has been chosen for the actual *management* of environmental quality from *stationary* sources is the states.

(5) Regulatory power without significant economic incentives is the technique chosen for actual abatement of harmful discharges.

(6) Municipalities are encouraged to abate through subsidization of waste treatment plants.

Mr. Kneese's recommendations of the best strategy for *managing* water and air pollution abatement is in marked contrast to existing public policy. Mr. Kneese argues convincingly for problem-shed management (e.g., river basins) and application of charges for using air and water common property resources (e.g., effluent fees). The national strategy is to rely on state governments for management. It generally ignores use of effluent or emission charges.

While Mr. Kneese identifies the asymmetry between the bargaining position of industry which must pay for pollution abatement and those who benefit, he fails to consider the asymmetry that exists between use of imperfect existing institutions and perfect new institutions. But he does observe the results of this phenomenon when he notes that river basin organizations have shown the ability to plan supposedly efficient and effective water abatement strategies but lack the financial or legislative support to actually manage water quality. Existing institutions with their associated beneficiaries and bureaucracies are jealous of their prerogatives and will resist new institutions that would share political and social and economic power. Existing institutions obviously have the advantage over any proposed institution with less certain benefits and beneficiaries and no existing bureaucracy to assure its birth and survival.

Usually compromises occur and only second-best institutional arrangements are feasible. The extent that the compromise is efficient can be influenced by the analysis of alternative institutions and the value of shifting to them. The value of a change in institutions must be significant to make it worth the cost of the change. This may mean that the cost of water quality management by states must be greater than, say, 25 percent before it is worth abandoning it for river basin management, for example.

This criticism of Mr. Kneese's inadequate treatment of institutions is not intended to imply that the cost differences of alternative institutions is not great enough to justify major changes. I believe they are. Nor is it in-

tended to diminish the importance of considering all factors that affect the costliness of environmental quality management. Mr. Kneese recognizes this in his discussion of the studies of the Delaware River Basin.

In the case of air pollution abatement, now that the Federal Government has made a strong commitment to establish national ambient, air-quality standards and to insure that they are met, it is very important to consider least costly strategies and the associated institutions which can achieve these standards. In preparation for the policy decision leading to the President's proposal of the Clean Air Act of 1967, we attempted to crudely identify the variations in costs from alternative abatement strategies, which themselves are dependent on variations in institutional arrangements. I was frankly surprised to find the lack of information which, disappointingly, has not improved much since then. We didn't know what damage to human health occurred from air pollution, except from "structured intuition" and some fragmented data. Only the roughest indication of the damage to physical objects was available. Accepting the fact that air quality standards will be based on such scarce data and judgments (including the inherent redistributive nature of pollution abatement policies), we turned to measuring the cost of achieving given levels of abatement by alternative strategies.

Lacking anything better, we used an air emissions model of New York and added some industries not found in New York so as to have a more "typical" model of the "average" air pollution problem in the nation. Lacking a meteorological or diffusion model for New York, we borrowed the only diffusion model available, which was for St. Louis, and proportionally shrunk the New York emission model to fit a population of two million. We heroically assumed that the resulting model was typical of all metropolitan areas (SMSA's) and that the pollutants we modeled—sulfur oxides and particulates—were a problem only in metropolitan areas and could be considered solely by concentration levels on human beings. We gathered the available data on the least costly ways the industries, space-heating plants, and households in New York could reduce SO_x and particulate emis-

sions. For example, fuel substitution (e.g., #2 oil for residual oil), processing of inputs (e.g., desulfurization of oil), process changes, and treatment of pollutants at the emission point (e.g., scrubbing and bagging). Based on an objective to reduce human exposure to sulfur oxides and particulates, we programmed alternative abatement strategies.

First, a strategy was developed to guarantee that SO_x and particulates would not exceed a given level of human exposure. The given level was roughly the *annual* average of 60–75% abatement of SO_x and particulates levels. Abatement of the pollutants occurred only when this level would be exceeded, including abatement, when feasible, for as short as a day. When the “typical city” was well ventilated, no abatement was required even though heavy discharges were occurring because the concentrations on people didn’t rise to the threshold abatement level. The relative accuracy in forecasting temperature inversion, calm wind days, and major shifts in wind patterns was important. The additional annual cost estimated for this abatement strategy was about \$250–400 million.

Second, an abatement strategy was followed that would reduce SO_x and particulate concentration on human beings by 60–75% from what would have been the level without abatement for each day of the year. Abatement was accomplished by requiring only those emitters to abate that caused an increase in pollutant concentrations on people and those emitters that could do it at the lowest cost. The results indicated an additional annual cost of about \$750 million, about half to industry and half to households. Industry’s costs would increase value-added by about one-sixth of one percent. Household costs would increase by about \$3 per person. Surprisingly, capital costs were only 15% of *direct* abatement expenditures. Operating costs, for example, from higher fuel costs associated with the substitution of low sulfur fuels, accounted for 85% of the additional annual costs. Under these conditions, capital subsidies like tax credits or rapid depreciation allowances would be of little assistance to industry.

Third, a *uniform* abatement strategy was tested where all polluters within the metropol-

itan area were asked to reduce SO_x and particulates proportionally until the goal of 60–75% reduction in pollution levels was achieved. The additional annual cost was estimated to be about \$1.3 billion. Interestingly, industry’s abatement share increased more than proportionally from the strategy described in strategy two, above. This occurred because industry discharge is often located downwind of the concentration of people as compared with discharge from space heating plants of office buildings, apartments, and homes. This was also shown to be true in a separate study of electric powerplant locations and prevailing wind patterns in the 20 largest U.S. cities.

Fourth, all emissions within and outside metropolitan areas were required to abate proportionally to reach a 60–75% reduction of SO_x and particulate concentration on human beings. The additional annual cost was more than \$2 billion for this approach.

Although there is no claim for accuracy or scholarship, these estimates were felt to be within the range of minus 50% and plus 500% of the real cost, and this was all the accuracy that was needed to recognize the advantage of managing air common-property resources efficiently, including working for the creation of air-shed institutions.

I am firmly convinced that policymakers as well as industry and the public, who will ultimately pay the cost, are ignorant of the great differences in costs based on different abatement strategies. Without this information, the most easily understood and administered strategy would be selected which is to require uniform abatement for everybody. Equal pain is a tried-and-true administrative principle, but unfortunately the cost could be from 200% to 1,000% greater than by a highly-selective abatement strategy.

In regard to ignorance, Mr. Kneese attributes too much current rationality and knowledge to businessmen, conservationists, and government officials. In addition to the obvious incentives to delay abatement, I found key businessmen ignorant of the likely effects of Mr. Kneese’s recommendation of problem-shed management and effluent fees. They correctly saw the creation of another tax and further infringement on free enterprise, but they failed to realize that Mr. Kneese’s approach

would mean a smaller government bureaucracy hounding them into compliance and a smaller cost than is likely to occur with the regulatory approach that appears most popular.

Conservationists generally felt that effluent fees are "a license to pollute." They couldn't understand the incentive effect that effluent fees would have on reducing harmful discharges at the lowest cost.

Municipalities could only see the effluent fee as an additional tax on their already declining and regressive tax base. Understandably, they were quick to recognize the revenue-sharing nature of subsidies for waste treatment facilities that would have been built in most cases anyway and the assistance this would have to ease their revenue problems.

Mr. Kneese failed to recognize that we do not have the technology to measure discharges easily or inexpensively so as to monitor an effluent or especially an emission charge system. However, given the advantages of such an approach, improvements could be easily stimulated and be available within a few years.

On efficiency grounds, I find myself in agreement with Mr. Kneese's preferred public policies for managing air and water common property resources. However, I feel that Mr. Kneese fails to consider feasible institutions and the transition to them. I would emphasize more forcefully that alternative strategies for management of air and water quality, with their associated institutions, can affect cost by not merely 25% or so, but by 100, 200, or even 1,000%. The major policy issue for the next few years is to deal with this problem.

RICHARD W. JUDY: Dr. Kneese has provided us with an overview of the state of economic research and policy in the area of environmental quality management. He argues persuasively that economic theory has until recently contributed little of direct value to policy formation and management in this area. The theory of external effects is not sufficiently robust to provide much insight into pervasive externalities. Welfare economics generally washes its hands of externalities by assuming that bargaining or appropriate tax-subsidy adjustments can correct matters.

At the root of environmental pollution, as Kneese argues eloquently, is the fact that certain commonly owned assets with positive marginal opportunity costs carry a marginal user charge of zero. This, in turn, springs from the absence in our system of institutions charged with exercising ownership responsibilities over these common property resources. As Kneese puts it: "... now that many properties of the natural environment have clearly become scarce resources, we are confronted with a vast asymmetry in the ability of our property institutions to form the basis for efficient resources allocation." We need to recognize that our environment is capable of providing many different services, some of which are mutually incompatible. We need mechanisms and institutions that will allocate these many uses among many users efficiently and within the bounds of generally accepted standards of equity. We are very far from having those mechanisms and institutions in operation although some of us think we know something about how they should operate.

The traditional institutions for dealing with pollution are two-pronged: statutes forbidding discharge of wastes into the environment, and public subsidies for the construction of treatment facilities. Together, these two policy instruments are expensive and ineffective. From studies of the Delaware Estuary and other water basins, we know that various alternatives to waste treatment exist and that they often offer less expensive routes to desired water quality. Low-flow augmentation and in-stream reaeration are two examples of alternative, and often superior, collective pollution control measures. At the point of waste generation, there are numerous possibilities for altering product and input mixes, changing processes, recycling wastes and otherwise diminishing the quantity and noxiousness of residual wastes discharged into the environment. Use of these alternatives in water quality management programs, Kneese argues, would bring major gains in efficiency. But, he says, our existing collective action institutions are ill-suited to realize these gains.

The diagnosis, then, is inadequate incentives and institutions. Kneese prescribes two measures: (1) the creation of "problem shed"

management agencies and (2) the use of effluent charges. Many of us in Canada have followed a similar path of reasoning and you may be interested to know how we are progressing in our efforts to create more adequate institutions and incentives.

First, let me say that we, just as you, are bedevilled by constitutional ambiguities and constraints. By our Constitution, the ownership of natural resources lies with the provinces. The federal role is confined to international and boundary waters or elsewhere where the "national interest" is at stake. It is ironic that while our Constitution theoretically grants wide powers to the Federal Government to ensure "peace, order, and good government," in fact, our provinces are much stronger vis-a-vis Ottawa than your states are vis-a-vis Washington. This obviously constrains the federal role but, because of ambiguous assignments of specific powers, no one is really sure just where the constraints lie.

In any case, a recent federal reorganization has thrown all agencies with environmental management responsibilities into a new Department of the Environment and Renewable Natural Resources. At present, the department is headed by the former Minister of Forestry and Fisheries, the Hon. Jack Davis. The Federal Parliament recently enacted the Canada Water Act. This legislation provides for two key activities: (1) the preparation of comprehensive river basin plans by jointly funded (50-50) and staffed federal-provincial study groups, and (2) the establishment of river basin management agencies jointly by the federal and provincial governments. These agencies normally will report to both federal and provincial elected authorities and derive funding from both levels in proportions that are unspecified by the Act.

The legislation specifically provides for the river basin management agencies to construct and operate a wide variety of collective measures including low-flow augmentation, diversion, reaeration and treatment facilities. It also specifically provides for effluent charges and user fees as well as for the establishment of ambient water quality standards. Credit for this excellent act belongs largely to Dr. E. Roy Tinney and Assistant Deputy Minister

Al Davidson, both of the Department of Environment and Renewable Natural Resources.

Comprehensive river basin plans are now under way in four basins and they involve five provincial and the federal governments. Many more will begin soon. The time to complete the studies will vary from one to three years. As you would expect, many things remain to be worked out. How should we go about establishing the ambient water quality standards? How do we evaluate the magnitude and incidence of environmental benefits and costs that stem from various features of the plans? Should there be public participation in establishing the objectives and evaluating the alternatives? If so, how and when should it happen? Should we use the effluent charge scheme or some other economic instrument such as the discharge warrant scheme proposed by my colleague, Professor John Dales? If we do use it, how should the charges be set? If certain key federal civil servants and their consultants think that one of these schemes should be employed, how do we persuade other important but dubious federal and provincial officials? There are so few actual examples of effluent charges at work and they are very shopworn. This, I think, is where the Russell-Spofford Model (or one similar to it) may serve a useful purpose if some empirical flesh can be put onto its bare mathematical bones; it can be the vehicle for demonstrating *with numbers* the points that noneconomists frequently do not grasp from abstract economic reasoning alone.

Perhaps most fundamentally, some of us who are involved in this exercise in "institution building" are asking, "What should be the objectives of the new water quality management agencies? What should be the explicit criteria by which their performance is judged and their officials rewarded?"

To many of these questions there exist no neat answers. They must be worked out in the process of implementing the legislation. But for many of the answers that we do have, we are greatly indebted to Dr. Kneese and his colleagues at Resources for the future.

MARC J. ROBERTS: Let me begin with a traditional disclaimer. I find myself in sympathy

with much of what Dr. Kneese has said. There are places, however, where his treatment seems unclear, or his suggestions subject to limitations that he has not fully spelled out.

Why should private bargaining in some cases not produce a Pareto-optimal level of output for external effects? That is, why should unexploited gains from trade remain? (In what follows, no distinction is made between "public goods" and "externalities" since recent work has clarified that these are analytically indistinguishable.) In disagreement with Kneese, "an unequal distribution of economic power" does not seem to be one such factor. Unequal power could plausibly affect the division of nonmarginal gains among the parties. But why should it lead the bargainers to choose anything but the largest pie to cut up? There are, however, five distinguishable elements to a situation that might lead to non-optimal results.

(1) *Bargaining and decision-making costs:* The higher such costs, the more likely that the effort required to reach a solution will not be—or appear not to be—worthwhile. These costs increase with the number and heterogeneity of the participants in the bargaining, and with the complexity of the technical options for dealing with the externality. On both counts, water pollution control is a poor candidate for private action, especially since, as Kneese suggests, large scale capital investment and complicated river management systems are often required.

(2) *Difficulties raised by the nature and distribution of benefits:* What happens to private bargaining when the benefits of such agreements are small for many potential participants, or are difficult to measure, or depend upon the development of (or changes in) the tastes of various actors? Negotiation and bargaining are more likely to appear to be unattractive in such cases—especially if these are high-cost activities. Again, water quality is a perfect example. How much would you pay to have the probability of a migratory fish kill on the Delaware River changed from .09 to .08? How much more will people pay for swimming facilities after they are available and they have learned to swim than they would be willing to pay *ex ante*?

(3) *Non-excludability and the "free rider" problem:* If large numbers are involved, and if all of us cannot prevent you from enjoying the benefits of clean water, then you have no incentive either to reveal your preferences or to pay the costs of becoming involved in negotiations. Even where exclusion is technologically possible, it will be statically inefficient if consumption is nonrival in the relevant range. This is often the case in water quality—e.g., we can charge for swimming and boating. At the same time, many important aesthetic benefits are not marketable. The classic dilemma of nonrevelation of preferences, however, has perhaps been overdone. In presenting his "public goods" model, Samuelson assumed that each level of expenditure on a public good-externality provided a given distribution of utility. Since man X could not be deprived of the good without depriving everyone else, man X could safely misstate his preferences. If public decisions were based on the utility distribution, he had every incentive to do so. However, as Buchanan has recently shown, more typically the distribution of benefits from publicly provided goods and services can be varied. A given expenditure level maps to a frontier, not to a point, in individual utility space. We can clean up some parts of a river more and other less, or vice versa. The threat of being at least partially deprived helps to explain why nonrevelation (in the game-theoretic sense) does not seem to be much of a problem in water quality management. (Another possible explanation is that the society does not worry about the distribution of subjective utility, but, instead, maximizes an objective function defined over income and a vector of selected consumption opportunities such as health, housing, education, etc.)

(4) *Irrational or strategic moves:* In small numbers cases especially, bargainers may threaten moves that are apparently irrational (i.e., produce non-Pareto optimal outcomes) as a way of trying to obtain a larger share of the infra-marginal gains. Such threats become more likely when the players are involved in a whole series of bargaining situations and overall credibility really matters—e.g., industries and local governments in a given area.

(5) *Legal limitations:* The definition of property rights and liability rules can affect

the outcome. The definition of these rights can affect the level and distribution of bargaining costs or even make bargaining impossible. Further, different definitions of property can affect the outcome through their impact on total conditions. The maximum profit output of a paper mill that is optimally paid by sports fishermen to clean up some of its waste may be the same as its output when it optimally pays the fisherman for the right to dispose of waste in their river. But in one case the profit can be greater than zero and in the other less.

This review simply reinforces Kneese's conclusion. Private bargaining is seldom a viable option for dealing with environmental quality externalities—as the lack of such bargaining attests. Government action is designed exactly to cut through this tangle. Collective provision can result in lower transaction costs, can alter the distribution of infra-marginal gains and can change the pattern of incentives for preference revelation. At the same time, public provision, like market provision is often (always?) imperfect. The archtypical policy problem is to choose among a variety of imperfect arrangements in a specific case. Theoretical arguments alone are insufficient. Every economist has his favorite example of government regulatory actions that have led to results less desirable than the externalities they were designed to correct.

To move on, what can we make of the Russell-Spofford model that Kneese reviews? It is nonstochastic, nondynamic, and assumes we know all the relevant production functions. It is meant to be solved for the optimal solution. I agree that large-scale simulation models are invaluable tools for studying the behavior of systems too complicated to think about unaided. But let us not mistake a heuristic aid for anything more. The aspects of the situation that are left out of such models often are crucial in real life. Perhaps that is why actual decisions are so often incremental, iterative and sequential—as opposed to globally optimizing in character.

The real policy problem is both dynamic and stochastic. For example, the production functions for waste output are in a period of rapid change, as it becomes valuable for managers to begin to worry about waste disposal. In addition the size and composition of pro-

duction, the distribution of population, land development and drainage patterns, agricultural and mining practices all will continue to change. Thus the "opportunity set" is rapidly and continually altering. Benefits too will continue to change, as tastes, income, experience, and education all continue to vary. Such developments limit the value of an analysis based on current experience. A more relevant model would, for example, need to take into account the costs and virtues of incremental policy steps designed to minimize the social risks of mistaken forecasts.

There are still other problems. No allowance is made for the second-best character of the real world. The interlinked problems of investment timing, irreversible effects, and option demand are omitted. The allocation of risks is also not illuminated. How can it be in a nonstochastic formulation? No mention is even made of income distribution! More technically, the damage functions for various residuals are apparently assumed to be independent—contrary to actual fact. The moral is that, however complex and ingenious the model at hand, it is too imperfect to be more than a helpful heuristic device.

One additional limitation deserves mention, in part because it is characteristic of so much water resource planning. The economic and political processes that link water resource decisions to the lives of individual citizens are largely omitted. As far as I can tell from Kneese's brief treatment, final goods prices and tax payments (direct and via price effects) are all left out. No wonder income distribution is not raised as a criterion. Adequate systems analysis does depend upon appropriate definition of the system.

Another question which bears both on this model and on Kneese's policy recommendations is the notion that air, water, and solid waste planning (and administration) should be carried out in a unified manner. This contention is not self-evident. On the contrary, the sphere of analysis or administration should be determined not by whether there are interactions, but by whether these are strong enough to warrant attention. Failure to pose the question this way would lead to a single worldwide model (or bureaucracy) for everything.

Innumerable economic and social activities interact with water quality. Since water quality depends upon quantity, optimal pollution control is related to all flow regulation functions: hydroelectric power, navigation, and flood control. Diversions from the stream affect quantity and quality, and policy here is, in turn, interdependent with inter-basin water transfers and ground water management. Many types of activity are relevant because they affect the waste load; industrial and residential development, agricultural and mining practices, land drainage, and even street cleaning. Further, the magnitude of the benefits depends upon the recreation and fishery facilities provided. These various directly water-related policy problems in turn cannot be solved without taking into account transportation, power, and recreation planning in general; all locational incentives and regional development programs, etc.

Why then include air and water pollution in a single model, as opposed to, say, water quality and water supply? In fact, the latter linkage is more important in many parts of the country. The scope of the Russell-Spofford model thus needs more justification than we have been given. Indeed, it is to be expected that models with differing orientations will be more or less helpful in different places, depending upon which interfaces are most important in a given area.

Now for some specific policy issues, Knease notes several reasons for having a basin agency for water quality, and, as I have indicated elsewhere, in general I strongly agree. I would contend, however, that the basin agency should have a monopoly of waste treatment plant construction and operation. If we allow intermediate "public utility" treatment firms to arise, their charges to waste sources would have to be regulated in complex detail. Marginal charges along several waste dimensions, as well as total charges, would need to be controlled. This is an unpromising undertaking. It seems expensive and unreliable to try to achieve short-run variations in treatment plant operation (in response to stream conditions) by altering effluent charges. Rather, this seems exactly the sort of instance Coase had in mind in explaining the

origins of the firm—one where it is cheaper to achieve coordination via direct bureaucratic integration across market lines. In sum, decentralized approaches have no advantages over direct operation, and many deficiencies.

A comprehensive basin authority with sole right to dispose of waste in the stream would eliminate any role for pure effluent charges *per se*. Instead, *firms* would pay service charges for waste abatement, based on the marginal costs to the agency of maintaining stream quality standards set by the states and the federal government. Such charges would provide appropriate incentives for internal waste-reducing process changes. They would seem to be more politically defensible than effluent charge schemes. Setting such prices would be complex—involving administrative cost problems, marginal and total conditions, second best problems, and dynamic considerations. I will not try to explore their character further here.

Municipalities (and citizens), I suggest, should not pay charges (service or effluent) for the treatment or discharge of domestic wastes. Such charges would be close to head taxes and extremely regressive. If property taxes were used to raise this revenue, the burden would also be extremely regressive. Such fees have a dubious allocational rationale. Monitoring individual household waste output is much too expensive to be feasible. The few opportunities opened to households for altering their volume of sewered wastes involve the use of physical devices like garbage grinders. It seems more practical and equitable to tax these capital goods directly to achieve allocational goals. In a pure effluent charge situation, I am dubious of the value and equity of monetary penalties designed to coerce municipal treatment plant construction.

If treatment for domestic wastes is tax-supported, federal sources should be used to avoid a regressive impact. Avoiding regressivity here is important since the financing of most other environmental improvements (e.g., automobile emission control, cleaner electricity generation) will be quite regressive. Given the higher willingness to pay for environmental quality by the more well-to-do, federal financing is required here if we are to avoid

having the poor subsidize the rich in this area. Proponents of revenue sharing, who argue that funds are relatively more scarce at the state and local level, should be sympathetic to this proposal. The National Conference of Mayors and the National League of Cities have recently supported a basin authority water quality system. This indicates their willingness to give up bureaucratic independence in order to avoid the financing burden, operating problems, and possible enforcement actions associated with sewage treatment.

Having mentioned various specific points, it may be helpful to conclude by suggesting some of the fundamental questions raised by the discussion. First, let me concur in Kneese's suggestion that "institution building" is essential for environmental quality improvement. Several technical options will be difficult or impossible to use without a new organizational framework. Yet in these areas, purely economic considerations are insufficient. The presumption Bergson suggested more than 30 years ago—that economic welfare was separable from general welfare—is clearly violated here. Various administrative structures will have an impact on political and social values, on individual attitudes toward and involvement in political process, etc. Multi-disciplinary political economy is required in order to propose sound recommendations.

Next, the whole question of which units of government are competent to make which choices must be squarely faced. As long as income transfer systems are national and uniform (e.g., the income tax), local decisions on water quality and its financing will affect the

distribution of utility. Yet is this distributional question more properly a "national" one? Further, do we want to allow regions to opt for lower environmental quality and lower tourism, but higher industrial development? How are interregional effects and extraregional interests to be balanced against the values of decentralization and local choice?

Third, the limitations on our benefit measures are all too apparent. How can we develop more meaningful data? What can we do about endogenous preferences, an empirically important phenomenon in this area? Is society (or should society be) in the business of choosing a distribution of individual, idiosyncratic, subjective happiness? Are we prepared to face the ethical, conceptual, and informational difficulties taking this seriously would entail? In a similar vein, a more explicit treatment of the allocation of risks needs to be developed. In what cases are the private attitudes of beneficiaries and cost bearers sufficient for aggregation over states-of-the-world, and in what cases are there distinct social risks?

Finally, let me close by reiterating the importance of distributional issues. As long as the marginal social value of income to various citizens is not the same, acting as if it were (i.e., worrying only about "efficiency") represents a restrictive and dubious ethical presumption. The crucial role of the professional in my view is to clarify and make vivid the ethical implications of alternative choices. If, instead, he suppresses or quietly ignores these issues, he will only lead others to do likewise.

COMMERCIAL POLICY AND LESS DEVELOPED COUNTRIES

Trade Policies in Developing Countries

By BELA BALASSA*
Johns Hopkins University

The purpose of this paper is to examine some of the implications of the author's findings concerning the effects of trade policies followed by developing countries on their export performance and economic growth. In the discussion, use will be made of the results of several studies carried out in recent years: an investigation of the growth strategies of six semi-industrial countries, two of which, Argentina and Chile, are customarily included in the developing country category;¹ a comparative study of the structure of protection in Brazil, Chile, Mexico, Western Malaysia, Pakistan, the Philippines, and Norway;² and a study of industrial policies in Taiwan and Korea.³ Following a brief

summary of the findings, guidelines will be suggested for trade policies by developing countries.

I

Instruments of trade policy employed by developing countries include import tariffs and surcharges, export taxes and subsidies, multiple exchange rates, as well as quotas and licenses. By affecting the relative prices of inputs and outputs, these measures influence the allocation of resources, including new investment, and provide incentives—or disincentives—to import-substituting and export activities.

Discrimination among economic activities introduced by the use of trade measures exists to varying degrees in the nine countries studied.⁴ Argentina, Brazil, Chile, Pakistan, and—to a lesser extent—the Philippines provide considerable incentives to manufacturing industries at the expense of primary activities. The extent of discrimination in favor of manufacturing and against primary production is relatively small in Mexico; it is even less in Korea and Taiwan; and, on the average, virtually nil in Western Malaysia (for short, Malaya).

Discrimination in favor of manufacturing entails the protection of domestic manufacturing industries against imports

* This paper was written as part of a consultant arrangement with the World Bank.

¹ Bela Balassa, "Growth Strategies in Semi-Industrial Countries" *Quarterly Journal of Economics*, February 1970, pp. 24–47.

² Bela Balassa and Associates, *The Structure of Protection in Developing Countries*, Baltimore, Md., The Johns Hopkins University Press, 1971. For a discussion of some of the methodological issues and a summary of the estimates of effective rates of protection, see "Effective Protection in Developing Countries" in *Trade, Balance of Payments and Growth: Papers in International Economics in Honor of Charles P. Kindleberger* (J. Bhagwati, R. Jones, R. A. Mundell, and V. Vanek, eds.), Amsterdam, North Holland Publishing Co., 1971.—In the present paper, results pertaining to Norway will not be considered. Furthermore, in the case of Argentina, Brazil, and Chile, the paper will consider the situation existing prior to the changes in policies introduced in the late sixties. These policy changes will be briefly referred to at a later point in the paper.

³ Bela Balassa, "Industrial Policies in Taiwan and Korean," *Weltwirtschaftliches Archiv*, Band 105, Heft 1, 1971. To be reprinted in a volume in honor of Raul Prebisch.

⁴ The results cited in the paper have been taken from the four studies cited above and do not include recent changes in the policies that will be noted below. Some of the estimates on exports and growth have been reproduced in Tables 1 and 2.

whereas primary exports are penalized by tariffs on their inputs and by the lower exchange rate associated with protection. The protection of manufactured goods and discrimination against primary exports usually go hand-in-hand, the principal exception being Malaya, which levies a tax on its major primary exports, rubber and tin, but does not protect manufactures.

In Argentina, Brazil, Chile, and the Philippines, there is also a substantial bias in favor of import substitution and against exports in protected manufacturing industries. Firms producing for home markets can get the domestic price raised by high import protection while, in the absence of export subsidies, they could obtain only the world market price in exporting. By contrast, the bias against exports in manufacturing industries is relatively small in Mexico and Malaya where levels of protection are low and it is practically nonexistent in Pakistan, Taiwan, and Korea. In Pakistan, manufacturing industries receive high protection irrespective of whether they produce for domestic or foreign markets; in Taiwan and Korea the extent of import protection is rather low and, on the average, it is matched by subsidies to manufactured exports.

II

The trade policies followed during the postwar period have affected export performance and economic growth in the countries under consideration. The unfavorable treatment of primary exports has contributed to the decline in the shares of Argentina, Brazil, Chile, and Pakistan and, to a lesser extent, the Philippines in the world market for their major export commodities and has retarded the development of new exports. Malaya, too, has experienced a decline in the market shares of its major primary exports, but it has expanded its minor exports which receive more favorable treatment. Finally, Mex-

ico and to an even greater extent Taiwan and Korea have been successful in raising the world market share of their major primary exports and in introducing new export products. In the latter two countries, these developments have taken place during the sixties following a shift in economic policies from import substitution to export promotion.

The expansion of primary exports in individual countries has further been affected by world demand conditions in the markets for their major export commodities. The main beneficiaries of favorable world market trends have been Chile (copper) and the Philippines (oilseeds) while the slow growth of world demand and unfavorable price changes for rubber and tin have depressed export earnings in Malaya.

The trade policies followed, together with changes in world market conditions, largely explain intercountry differences in the rate of growth of primary exports. Table 1 shows Korea and Taiwan in the lead, followed by Chile, the Philippines, Mexico, Argentina, Malaya, Brazil, and Pakistan. Comparisons of growth rates are less meaningful for manufactured exports because several of the countries under consideration started from a very small base. Instead, we use the share of manufactured exports in manufacturing output and in total exports as indicators of success in exporting manufactures.

In countries with a substantial bias against the exports of manufactured goods, these commodities continue to account for less than 3 percent of manufactured output and account for at most 10 percent of total exports. Among countries with a lesser bias, Mexico's manufactured exports have reached 5 percent of manufacturing output and 25 percent of total exports. In Malaya 10 percent of manufactured output is exported although, given the relatively low share of manufacturing and the high share of exports in GNP, manufactured

TABLE 1—EXPORTS IN SELECTED DEVELOPING COUNTRIES

	Argentina	Brazil	Chile	Mexico	Korea	Malaya	Pakistan	Philippines	Taiwan
Average annual rate of growth of exports									
Primary goods									
1950-60	0.1	-0.9	5.3	4.0	-3.5	1.9	-7.5	5.4	-1.2
1960-69	3.6	4.2	8.4	5.5	21.0	0.6	2.5	6.0	16.4
1950-69	1.7	0.6	6.8	4.7	16.7	1.1	-2.9	5.7	8.3
Manufactures									
1950-60	-7.6	15.6	11.1	12.2	0.2	29.0	35.0	4.7	30.5
1960-69	17.3	19.1	10.7	19.9	69.0	12.7	14.5	25.0	34.0
1950-69	3.5	16.3	10.9	15.8	35.6	19.0	24.0	13.3	32.5
All commodities									
1950-60	-0.5	-0.6	5.6	5.0	-2.8	2.0	-2.3	5.4	3.6
1960-69	4.6	5.1	8.5	7.2	38.9	1.2	6.3	6.8	24.0
1950-69	1.9	1.1	7.0	6.0	18.2	1.7	0.4	6.0	14.9
Manufactured exports as a percentage of output—1969									
total exports—1969	2	1	3	5	18	10	8	3	36
	10	9	6	25	76	10	51	10	67

Note: For Taiwan and Korea the base year is 1953 instead of 1950. For Brazil, Chile, Malaya, Pakistan, and the Philippines, the terminal year is 1968 instead of 1969.

Source: National and international trade statistics.

goods provide no more than one-tenth of total exports.

In Taiwan and Korea, manufactured exports have been stimulated to a considerable extent by the adoption of export-oriented policies around 1960. As a result, exports of manufactures have increased to a considerable extent in the two countries, both as a proportion of total exports and of manufacturing output. By 1969, these proportions reached 67 and 37 percent in Taiwan and 76 and 18 percent in Korea.

Following the introduction of the Export Bonus scheme in 1959, manufactured exports have also assumed importance in Pakistan, with jute and cotton textiles being the principal items. But, owing to policies penalizing primary exports, foreign sales of raw jute and cotton have declined in an amount exceeding the rise in textile exports. Moreover, high subsidies to manufactured exports, together with the high protection of import substitutes,

have imposed a substantial cost on the national economy.

Owing to their favorable performance in both primary and manufactured exports, Korea and Taiwan are far ahead of the other countries studied in terms of the expansion of total exports (Table 1). Exports have increased more than the average also in Mexico where the growth of tourism and border trade (not included in the export figures) have further contributed to increases in foreign exchange earnings. For reasons noted earlier, exports have risen relatively rapidly also in Chile and the Philippines while increases have been small in Argentina, Brazil, and Pakistan.

III

The expansion of exports contributes to economic growth directly by raising national income and indirectly by providing foreign exchange for the import needs of

the domestic economy. An export-oriented policy also permits specialization according to comparative advantage—both between primary and manufactured activities and within the manufacturing sector. In particular, exports of manufactured goods enable firms to lower costs by employing large-scale production methods, reducing product variety, and participating in the international division of the production process through the manufacturing of parts and components for assembly abroad. Moreover, familiarity with foreign markets provides incentives for technological change and product improvement.

Import substitution, too, can be a source of economic growth in particular cases. A number of developing countries attained rapid rates of growth of manufacturing output and, to a lesser extent, national income, in the early stage of import substitution, which entails replacing the imports of nondurable consumer goods and their inputs by domestic production. Industries producing such commodities are the prime candidates for import substitution in developing countries since they employ chiefly unskilled and semi-skilled labor, do not require the application of sophisticated technology, and need few inputs from ancillary industries. Nor does the limited size of national markets constitute an important handicap for the development of these industries since the efficient scale of operations is relatively low and costs are not substantially higher in smaller plants.

But, in the absence of exports, the expansion of industries producing nondurable consumer goods and their inputs necessarily slows down after imports have been replaced since domestic production cannot continue to grow faster than home demand. Moreover, in the small domestic markets of developing countries, increasing difficulties are encountered in import

substitution in other intermediate products, capital goods, and durable consumer goods. These commodities have higher technological and skill requirements, require the availability of materials, parts, and components from other industries, and need large-scale production for efficient operations with costs being substantially higher at lower output levels. Last but not least, in the event of continuing protection, there will be few inducements for technological improvements.

These considerations help to explain intercountry differences in rates of economic growth (Table 2). In Taiwan and Korea, the growth of GDP has accelerated to a considerable extent following the adoption of export-oriented policies. In expanding the exports of nondurable consumer goods, the two countries have utilized their educated manpower while the capital requirements of these industries are relatively low. The leading role of exports in the growth process is indicated by the high incremental ratio of exports to GDP; in 1960–69, this ratio was 39 percent in Taiwan and 29 percent in Korea.⁵

Exports have also importantly contributed to Mexico's economic growth and the relatively low protection of manufacturing industries has limited the cost of import substitution. In turn, the low degree of discrimination among economic activities has made it possible for Malaya to attain a rate of growth of national income substantially above that for exports.

The remaining countries of the group are characterized by import substitution behind high protective barriers. In these countries, the relationship between import substitution and economic growth has been influenced by their market size and the level of their economic development.

⁵ The incremental exports-GDP ratio (the ratio of the absolute increase in the value of exports to that of GDP) was calculated in constant prices from data given in national and international sources.

TABLE 2—ECONOMIC GROWTH IN SELECTED DEVELOPING COUNTRIES

	Argentina	Brazil	Chile	Mexico	Korea	Malaya	Pakistan	Philippines	Taiwan
Average annual rate of growth of value added									
Agriculture									
1950-60	2.3	4.7	1.2	5.4	2.3	3.2	1.4	5.1	3.9
1960-69	2.0	4.2	2.5	4.0	4.6	3.9	3.7	4.6	5.0
1950-69	2.1	4.5	1.8	4.5	3.6	3.6	2.5	4.9	4.5
Manufacturing									
1950-60	4.7	8.8	3.3	8.0	13.6	5.1	7.8	10.2	10.1
1960-69	4.6	5.9	5.9	2.0	16.0	11.7	8.6	4.5	16.1
1950-69	4.6	7.5	4.5	8.4	15.0	8.0	8.2	7.8	13.5
GDP									
1950-60	3.4	5.8	3.7	5.8	5.0	4.1	2.5	6.8	6.9
1960-69	3.4	4.3	4.5	7.1	9.2	5.7	5.6	5.1	9.9
1950-69	3.4	5.1	4.0	6.4	7.4	4.8	4.0	6.1	8.6
Per capita GDP									
1950-60	1.4	2.8	1.2	2.8	3.0	1.1	0.4	3.5	3.1
1960-69	1.8	1.3	2.3	3.6	6.4	2.5	2.9	1.6	6.6
1950-69	1.6	2.1	1.7	3.2	4.9	1.7	1.6	2.7	5.1
Population									
1950-60	2.0	3.0	2.3	2.9	2.0	2.9	2.1	3.2	3.6
1960-69	1.5	3.0	2.2	3.4	2.6	3.1	2.6	3.4	3.0
1950-69	1.8	3.0	2.2	3.1	2.3	3.0	2.3	3.3	3.3

Note: For Taiwan and Korea the base year is 1953 instead of 1950. For Brazil and Malaya, the terminal year is 1968, for the Philippines, 1967.

Source: National and international statistics.

Thus, the expansion of manufacturing output has slowed down to a considerable extent in the Philippines after the mid-fifties by which time the "easy" stage of import substitution had been largely completed. Despite rapid increases in exports due to favorable market conditions, there has been a decline in the rate of growth of GDP as well.

Argentina and Chile had replaced practically all nondurable consumer goods and their inputs before the period under consideration, and their small domestic markets have made the expansion of industries producing other intermediate products, capital goods, and durable consumer goods both difficult and costly. These countries have built up an industrial structure which entails the use of small-

scale and often outdated production methods, inadequate specialization, and the manufacturing of products of low quality.⁶ Discrimination against agriculture, associated with the high protection of manufacturing activities, has further hindered their economic growth and helps to explain why they have experienced the lowest growth rates among the nine countries studied.

Brazil, too, had completed the first stage of import substitution prior to the period under consideration. Its large domestic market, however, provided possibilities for the continued expansion of manufacturing during the fifties, mostly

⁶ For a detailed discussion, see Bela Balassa, "Growth Strategies in Semi-Industrial Countries," *op. cit.*, pp. 45-46.

in intermediate products, capital goods, and durable consumer goods. But, as the possibilities for import substitution have been increasingly exhausted, industrial expansion has slowed down in this country also.

Pakistan had practically no industry prior to independence and it was able to achieve rapid rates of economic growth by substituting domestic production for the imports of nondurable consumer goods and their inputs. Subsequently, the adoption of the Export Bonus scheme has contributed to its relatively rapid industrial expansion. Continuing discrimination against agriculture has however adversely affected the growth of the Pakistani economy. If national income is measured at world market prices rather than at the domestic prices distorted by protection, increases in per capita terms appear to have been small.⁷

IV

The experience of the countries under consideration suggests the conclusion that while the protection of the manufacturing sector may permit rapid growth at an early stage of import substitution, it will eventually have adverse consequences for economic growth. Discrimination among industries does not permit specialization according to comparative advantage; the high protection of domestic industry induces the establishment of high-cost import-substituting activities; and the bias against exports retards the development of manufactured exports. Finally, in the absence of foreign competition, there will be little incentive for technical progress in small protected domestic markets.

The increasing difficulties experienced by countries at higher stages of import substitution have recently led some gov-

ernments to reconsider their economic policies. In Argentina, the extent of discrimination against primary production and exports has been reduced and the protection of manufacturing industries has been moderated through a simultaneous devaluation and a lowering of tariffs. Manufactured goods also receive export subsidies in Argentina and such subsidies have assumed an important role in Brazil. Furthermore, in Chile, an effort has been made to lessen the degree of over-valuation of the currency while subsidies have been used to promote the exports of manufactured goods.

Efforts made to reform the structure of protection, however, have gone only part of the way and further progress is made difficult by resistance on the part of vested interests. Businessmen are opposed to changes in the status quo which ensures comfortable profits, and they demand continuing protection from foreign competition, whether this comes from the industrial countries or from developing nations as in the case of LAFTA. Additional problems are that transition to a more open economy would entail dislocation in particular industries and regions.

It appears, then, that once an industrial structure geared to import substitution has been established, change becomes increasingly difficult. This observation points to the need for making appropriate policy choices at the time when a country embarks on an industrialization program. In the following, guidelines will be suggested for trade policies by developing countries.⁸ Apart from their application to countries at an early stage of industrialization, the guidelines can provide a basis

⁷For such an adjustment, see Bela Balassa, *The Structure of Protection in Developing Countries*, Ch. 2.

⁸In formulating the guidelines, the author has drawn on the results of the studies referred to above, his experience in advising developing countries on trade policies, and the pertinent economic literature. Limitations of space have not permitted, however, the detailed consideration of particular issues.

for improvements in the policies of countries presently engaged in import substitution.

V

International trade theory tells us that small countries which do not affect the prices of their exports and imports will maximize welfare by specializing in accordance with price relations on the world market. Developing countries can generally take their import prices as given and they will not affect the prices of most of their manufactured exports either. This will not be the case, however, for traditional primary exports whenever increases in the country's exports lead to a fall in prices. For these exports, then, the relevant decision rule will involve equating marginal costs to marginal revenue from exports rather than to price. This can be accomplished by converting foreign exchange earnings from such exports at a less favorable exchange rate or—what amounts to the same—imposing an export tax on them. Export tax rates on individual commodities should be set by allowing for the elasticity of world demand, the country's share in world exports, and the possible reactions on the part of foreign competitors.⁹

The application of these measures would take account of market limitations for traditional primary exports, without unduly discouraging their production as has often been the case in the past. A further question is whether manufacturing industries should be favored over nontrad-

itional primary production, and if so, to what extent and by the use of what measures. In this connection, note should first be taken of arguments for infant industry protection, designed to compensate for assumed differences between social and private profitability. On the firm level, such differences may arise if the lack of credit facilities, the overestimation of risks, or simply the desire to exclude the possibility of bankruptcy provide disincentives for investment, although eventual cost reductions through the learning process or through increases in the scale of operations would make the investment socially desirable. Other instances are when some of the benefits of the pioneering firm's activities are enjoyed by others who utilize the know-how generated by the firm or hire away skilled labor and technicians it has trained.

It has often been said that infant industry arguments justify using production subsidies rather than tariffs since the latter limit the size of the domestic market by raising the price of the commodity in question. But while tariffs contribute to government revenue, subsidies represent a claim on this revenue. Budgetary reasons, then, may explain why developing countries use tariffs in preference to production subsidies. In fact, tariffs often account for a large part of government revenue in these countries, and their replacement by other forms of taxation may encounter practical difficulties.

The arguments for subsidies in preference to tariffs gain in force in cases when a particular distortion or cost disability needs to be corrected. This will be so if the cost of industrial labor to manufacturing enterprises exceeds its social cost in the form of the output foregone in primary activities from which labor is drawn. In some overpopulated countries, such a situation may exist on family-type farms where the contribution of the marginal

⁹ An extreme case is that of coffee where producing countries would be advised to set the export tax (differential exchange rate) at a level calculated to ensure that domestic supply be equal to quota allocations under the International Coffee Agreement. In this way, profits due to the price-raising effects of quotas in international markets accrue to the government and, rather than providing incentives to surplus production, the proceeds of an export tax can be transferred to other activities where higher returns are obtained.

worker is said to be less than his consumption. There is further the possibility that unemployment will persist at the existing wage rate which cannot be reduced lest it decline below a socially acceptable minimum. Mining industries, for example, utilize relatively little labor and countries relying on mineral exports may not be able to fully employ their labor force without providing special incentives for labor use.

In the cases described, the appropriate measure would be subsidizing the use of labor rather than imposing tariffs. Tariffs encourage the use of labor as well as capital in protected industries and they favor using labor in such industries in preference to other sectors of the national economy. Moreover, tariffs may provide incentives for the development of industries that would not be profitable under free trade even if wages were nil. In such instances, there is a trade-off between employment and growth since resources are channelled into industries with relatively high costs. Some of these industries may also have limited possibilities for improving productivity as is said to be the situation in the Indian cottage industry that receives considerable inducements. Finally, while employment-creating measures tend to improve the distribution of income, they may adversely affect savings and hence the prospects for future growth. The choice between employment and growth, then, becomes a choice between present and future employment.

Subsidizing labor use may take the form of taxing output and rebating the tax on the basis of the number of employees. This method would encourage the expansion of labor-intensive industries which use a developing country's abundant resource, labor, and would also provide incentives for employing labor-intensive production methods. However, there is no reason to restrict the application of this method to manufacturing industries, but

it should be extended to all sectors other than family farming.

Particular cost disabilities, or handicaps of manufacturing industries owing to inadequate overhead facilities, can also best be corrected by specific action rather than by protection. But again, the provision of such facilities should not be restricted to manufacturing. Thus, roads and electricity are needed for agricultural activities, just as an increase in the educational level of the labor force would contribute to the development of industry as well as to the modernization of agriculture.

VI

The question remains whether, apart from temporary protection on infant industry grounds and the correction of particular cost disabilities, manufacturing should receive preferential treatment. In support of this proposition, it has been adduced that productivity tends to rise more rapidly in manufacturing than in primary production and that the expansion of manufacturing industries provides indirect benefits by inducing investments in other branches of industry and improving the quality of the labor force.

The first claim holds true if we compare manufacturing with agricultural activities that employ traditional techniques, although modern advances in agriculture offer possibilities for improvements in productivity. Furthermore, linkages, among industries often favor the establishment of related branches of manufacturing, but one should not condone on this basis the establishment of inefficient industries which supply inputs to other industries at a high cost. There is finally some merit to the argument that manufacturing contributes to improvements in the quality of the labor force to a greater extent than does even modern agriculture.

From the point of view of long-term

policy making, further consideration should be given to possible future changes in the supply and demand of primary products. In some developing countries either the supply of primary commodities or foreign demand for them would eventually prove to be a limiting factor for the country's economic growth. In such a situation, the preferential treatment of manufacturing industry, where supply and demand limitations are negligible, would be warranted not only vis-à-vis traditional primary commodities, but also in comparison to the primary sector as a whole.

These considerations indicate the difficulties encountered in appraising the claims made for the superiority of manufacturing over primary production. The difficulties are compounded if we attempt to quantify these alleged advantages. Nevertheless, one may argue that manufacturing offers *some* advantages over primary production in the form of labor training and in encouraging the expansion of related industry that do not enter into the profit calculations of the firm but benefit the national economy. Moreover, manufacturing will improve the growth potential of the economy whenever supply or demand limitations would eventually impinge on primary activities.

There is some presumption, then, in favor of promoting manufacturing industry in developing countries. The word "promote" is used advisedly as it includes protection of production for domestic markets (import substitution) as well as assistance to firms exporting manufactured goods. Since, for reasons mentioned earlier, bias against manufactured exports entails an economic loss, equal incentives need to be provided to production for domestic and for foreign markets. This can be accomplished by granting a subsidy to the exports of manufactured goods at a rate equal to the tariff applied to the same commodity, or by using differential ex-

change rates for the manufacturing sector. Given the cost and uncertainties of entering foreign markets, it might even be desirable to provide additional incentives to exports of manufactured goods on a temporary basis.

Further questions are what are "reasonable" rates of tariffs and export subsidies and whether all manufactured goods should receive equal treatment. Assuming that particular measures are used to correct special cost disabilities and that the employment objectives are served by a direct or indirect subsidy to the use of labor, as a first approximation one may suggest providing effective protection at equal rates to all manufacturing activities that have passed the infant industry stage.¹⁰ In this way, one would apply the "market principle" in the sense that firms will be established that are profitable under such conditions and existing firms would have to improve their operations, change their product composition, or disappear altogether. At the same time, nonessential imports could be restricted by levying excise taxes that bear also on domestic production.

The choice of a "reasonable" rate of tariffs and subsidies for mature industries in the developing countries will depend on the particular circumstances of the situation and on the range of other policy measures available to a particular country. It may be suggested, however, that since most developing countries have small domestic markets, they should aim at eventually reducing the net effective protection

¹⁰ In "Decision Rules for Effective Protection in Developing Countries" (mimeo, November 1970), Trent Bertrand provides an elegant proof of the proposition that maximizing welfare subject to the constraint that a certain amount of value added is generated in the manufacturing sector involves equalizing effective rates of protection within this sector. In a more general model, the desired amount of value added in the manufacturing sector and the rate of effective protection of this sector would be jointly determined.

of manufacturing to levels observed in countries such as Denmark and Norway, i.e., to approximately 10 percent.

Exceptions to the proposed equality of effective rates may be made if there is evidence that profitability on the firm level greatly understates (or overstates) the contribution of a particular industry to the national economy. But such exceptions should apply to entire industries rather than to individual firms and only in cases that are well-documented so as to avoid a "slippage" in protection. In other words, the burden of proof should be on those who request favorable treatment.

Standard rates of protection should be applied also in the case of infant industries and one should avoid "tailor-made" tariffs. While it is difficult to judge how much protection would be justified on infant industry grounds, it does not appear likely that, exceptional cases aside, a rate of effective protection more than double that for mature industries would be warranted.¹¹ This additional protection of infant industries should be set on a declining scale so that its eventual disappearance provides incentives for improvements.

VII

The described scheme may be implemented by using a basic exchange rate for nontraditional primary products, export taxes on traditional primary exports, and a combination of tariffs and subsidies on manufactured goods. The same result could be achieved by applying differential exchange rates to the three groups of commodities, with further adjustments made

for differences in the elasticity of demand among traditional primary exports. The choice between the two alternatives, or a combination thereof, would have to be made on the basis of considerations of political and administrative feasibility, with further account taken of the implications of the choice of exchange rates for invisibles and for capital movements.¹²

Compared to the policies of industrial protection followed by developing countries engaged in import substitution, the application of these guidelines would entail providing more favorable treatment to nontraditional primary commodities, reducing the protection of manufactured products, and equalizing the incentives for manufactured goods sold in domestic and in export markets. Also, as a general rule, equal incentives would be provided to all branches of manufacturing other than infant industries, and additional protection to infant industries on a temporary basis.

For countries that have already embarked on industrialization behind high protective barriers, the application of the guidelines would entail a revamping of the structure of protection. Needless to say, this could not be undertaken instantaneously but would require a transitional period, the length of which would depend on the particular circumstances of the country in question. There would also be differences in the mode of application of these guidelines, again depending on political and institutional factors. Finally, the relative emphasis on direct measures and on the tariff-subsidy scheme may differ among countries at different levels of industrialization.

¹¹ According to an OECD study, economies of scale and external economies can hardly justify effective protection of infant industries exceeding 20 percent even if direct subsidies to labor use are not provided. This figure declines to 10 percent if labor use is subsidized. Cf. Ian Little, Tibor Scitovsky, and Maurice Scott, *Industry and Trade in Some Developing Countries*, London, Oxford University Press, 1970, pp. 158-59.

¹² In some cases, however, neither of these alternatives might be feasible because of constraints in policy making. In Chile, for example, the government is said to have obligated itself not to levy special taxes on copper, the major export commodity. Accordingly, the basic exchange rate would have to be applied to copper while tariff-subsidy schemes would need to be used both for nontraditional primary products and for manufactured goods.

Tariff Preferences and Separable Utility*

By CHRISTOPHER CLAGUE

University of Maryland

A prominent model of the trade diversion and trade creation resulting from tariff discrimination is the formulation by P. J. Verdoorn [11], referred to as the "Dutch" model by Harry Johnson in his footnote exposition of it [7]. The theory makes use of two elasticities—a share elasticity of substitution between partner and nonpartner imports, and an elasticity of import demand, applicable to all imports. The simpler version of the theory assumes infinite elasticities of export supply. The analysis is partial equilibrium in nature; trade in a particular category is analyzed and income and exchange rates are held constant.

This paper will discuss the implications of the Verdoorn-Johnson formula for the utility function of consumers. Specifically, if we add some not unreasonable assumptions, the Verdoorn-Johnson formula implies that the utility function is separable, with one of its branches homogeneous. These properties can be used to clarify the nature of the assumptions being made about consumer behavior. I will suggest that the failure to take these implications into consideration may have led to the overestimation of the trade diversion (as opposed to trade creation) resulting from tariff discrimination.

The present paper then proposes a slightly different form for the separable utility function to be used in the analysis of tariff discrimination. Supply elasticities

are also brought into the model. The relevant elasticities are then selected, either from the literature or by guesswork, and the model is used to predict the trade effects of a fifty percent tariff preference on manufactures and semimanufactures in favor of the less developed countries. The preference-granting countries are the United States, the United Kingdom, the Common Market countries, and Japan.

I

In the Verdoorn-Johnson model of customs union formation, the percentage increase in partner imports resulting from the discriminatory tariff reduction is made up of two components: (1) the substitution of partner for nonpartner imports, and (2) the partner's share in the increase in combined imports.

The equation is:

$$(1) \quad M_1 = (E_{s1}\pi + n_m P_m)$$

where

M_1 = partner imports in a particular category

π = the price of partner imports relative to nonpartner imports

P_m = the average price of combined imports (partner and nonpartner)

E_{s1} = the elasticity of the partner's share in combined imports with respect to π (a positive number)

n_m = the price elasticity of demand for imports (a positive number) and dots above the variables indicate proportional changes.

The effects of the union on nonpartner imports (M_2) are¹

¹ This derivation is from [1], which is the only clear

* Research supported by the General Research Board of the University of Maryland. Computer time provided by the University's Computer Science Center. The author would like to thank Clopper Almon and Lloyd Atkinson for very valuable comments, while absolving them of responsibility for any errors.

$$(2) \quad M_2 = (E_{s2}\dot{\pi} - n_m P_m)$$

where E_{s2} = the elasticity of the nonpartner's share in combined imports with respect to π (a positive number).² If export supply elasticities in partner and non-partner countries are infinite, the price changes in the formulas above can be calculated from tariff changes.

Equation (2) was employed by Lawrence Krause [8, pp. 50, 250-252] to estimate the effects of the Common Market on U. S. exports, on the assumption of infinite export supply elasticities. Krause, citing empirical work by Verdoorn, employed values of 0.5 and 2.0 for n_m and E_{s2} . Edwin Truman [10, p. 257] has pointed out that these values of the parameters in Johnson's formula imply that the formation of a customs union would result in trade diversion being larger than trade creation, as long as the common external tariff is an average of the partner countries' preunion tariffs, and as long as export supply curves are flat. Yet Truman and others [2, 4, 9] have found remarkably little evidence of trade diversion within the Common Market, although there has been a great deal of trade creation. Truman and Balassa attribute the small size of the observed trade diversion to the fact that the Common Market countries experienced rather full employment during the period covered by the empirical studies just cited, and consequently the export supply elas-

ticities of member countries were quite low. Without denying this interpretation, I would like to suggest an alternative—that E_{s2} has been assumed to be too high relative to n_m . My reasons for this statement will become apparent after the application of utility theory to the Verdoorn-Johnson formula has been explained.

I shall assume that we can treat a mass of consumers as if it possessed a single utility function with the usual properties. (The application of utility theory to customs unions has already been made by Janssen [6, esp. pp. 64-77, 140-143] who has also criticized Verdoorn's use of 0.5 and 2.0 as values of n_m and E_{s2} . All we are adding to his work is the employment of a separable utility function.) We shall also assume, for the purposes of this section, that the utility function consists of only three goods—partner imports (M_1), non-partner imports (M_2), and domestic production (M_3). All of these relate to a particular category of commodities, say shoes. (The possibility of the substitution of shoes for other commodities will be introduced in the next section.) Now consider the second term on the right-hand side of equation (1). P_m should be interpreted as the price index of imports relative to domestic goods. The assumption of a common elasticity of demand for the two types of imports implies that they must respond in the same way to a change in the price of domestic products. In other words, the cross-elasticity of demand for partner imports with respect to the price of domestic products (n_{13}) must be the same as the cross-elasticity of demand for nonpartner imports with respect to the price of domestic products (n_{23}).

We shall now add an assumption not strictly implied by the Verdoorn formula: We shall assume that the income elasticities of M_1 and M_2 are equal, or $n_{y1} = n_{y2}$. If we write down the Slutsky partition of the cross-elasticities of demand we have

exposition of Johnson's formula that I have seen.

$$M_2 = q_2 \cdot M$$

where q_2 is the share of M_2 in combined imports (M). We assume all prices are equal to 1 initially.

$$\dot{M}_2 = \dot{q}_2 + \dot{M}$$

where dots refer to proportional changes. By definition, $n_m = -\dot{M}/\dot{P}_m$ and $E_{s2} = \dot{q}_2/\dot{\pi}$. Equation (2) in the text follows immediately.

² Since $E_{s1} = -\dot{q}_1/\dot{\pi}$, where q_1 is the share of M_1 in combined imports, and since it can be shown that

$$-\dot{q}_1 = \dot{q}_2 (M_2/M_1)$$

it follows that $E_{s1} = E_{s2} (M_2/M_1)$

$$n_{13} = n_{13}^* + w_3 n_{y1} \quad \text{and} \quad n_{23} = n_{23}^* + w_3 n_{y2}$$

where n_{13}^* and n_{23}^* are the Slutsky cross-elasticities and w_3 is the budget share of commodity 3. It can easily be seen that our assumptions that $n_{13} = n_{23}$ and $n_{y1} = n_{y2}$ imply that $n_{13}^* = n_{23}^*$.

On these assumptions, we can show that the utility function can be written in a separable form:

$$U[M_a(M_1, M_2), M_3]$$

where the branch utility function M_a is homogeneous. We establish separability by using a slightly modified form of Goldman and Uzawa's Theorem 5 [5]. The theorem we need states that, in a well-behaved utility function, a sufficient condition for the separability of the utility function into two groups is that, for i and j belonging to one group, and k belonging to the other

$$\frac{K_{ik}}{K_{jk}} = \frac{\partial M_i / \partial Y}{\partial M_j / \partial Y}$$

where K_{ik} and K_{jk} are the Slutsky cross-partials, i.e., $K_{ik} = (\partial M_i / \partial P_k)^*$. In our case, since we have $n_{13}^* = n_{23}^*$ and $n_{y1} = n_{y2}$, it is easy to show that³

$$\frac{K_{13}}{K_{23}} = \frac{\partial M_1 / \partial Y}{\partial M_2 / \partial Y}.$$

Thus we have established separability.

(The reason for our slight modification of Goldman and Uzawa's theorem is that we are dealing with only two groups of commodities, while they restrict their theorems to cases of three or more groups. In the two-group case, "strong" separability and "weak" separability are indis-

tinguishable, and separability does not imply additivity.)

We establish the homogeneity of M_a by showing that, with constant P_1 and P_2 , the ratio M_1/M_2 remains constant as M_a changes. With P_1 and P_2 constant, M_a changes only as P_3 and Y change. But our assumptions that $n_{13} = n_{23}$ and $n_{y1} = n_{y2}$ imply that M_1/M_2 is unchanged by changes in either P_3 or Y . It follows that M_a is homogeneous in M_1 and M_2 .

Now we are in a position to appreciate the inappropriateness of the values 0.5 and 2.0 for n_m and E_{s1} . (Krause actually assumed a value of 2.0 for E_{s2} , not E_{s1} , but our argument comes to much the same thing in either case. See footnote 2 above.) First we must point out the relationship of n_m and E_{s1} to certain elasticities of substitution. Let us define:

s_{12} = the elasticity of substitution between M_1 and M_2

s_{a3} = the elasticity of substitution between imports (M_a) and domestic production (M_3)

q_1, q_2 = the shares of M_1 and M_2 in total imports ($q_1 + q_2 = 1$)

h_a, h_3 = the shares of M_a and M_3 in total consumption ($h_a + h_3 = 1$).

It can be shown that

$$(3)^4 \quad E_{s1} = q_2 s_{12}$$

$$(4)^5 \quad n_m = h_3 s_{a3} + h_a n_{y1}$$

⁴ By definition, $E_{s1} = -\dot{q}_1/\dot{\pi}$ and

$$s_{12} = -\frac{\Delta(M_1/M_2)}{M_1/M_2} \frac{1}{\pi}.$$

It can be shown that $\dot{q}_1 = \dot{M}_1 - \dot{M}$, where M refers to combined imports, and

$$q_2 \frac{\Delta(M_1/M_2)}{M_1/M_2} = \dot{M}_1 - \dot{M}.$$

It follows that $E_{s1} = q_2 s_{12}$.

⁵ The Slutsky equation for n_m is

$$n_m = n_m^* + h_a n_{y1}$$

To establish (4) in the text, it remains to show $n_m^* = h_3 s_{a3}$. If P_3 does not change, and all prices are 1 initially, then

³ $n_{13}^* = n_{23}^*$

$$\left(\frac{\partial M_1}{\partial P_3}\right)^* \frac{P_3}{M_1} = \left(\frac{\partial M_2}{\partial P_3}\right)^* \frac{P_3}{M_2}$$

$$K_{13} \frac{1}{M_1} = K_{23} \frac{1}{M_2}$$

$$\frac{K_{13}}{K_{23}} = \frac{M_1 \partial M_1 / \partial Y}{M_2 \partial M_1 / \partial Y} = \frac{\partial M_1 / \partial Y}{\partial M_2 / \partial Y}.$$

It is apparent that the values assumed for E_{s1} and n_m , together with the data on the market shares q_2 and h_3 , place certain restrictions on s_{12} , s_{a3} , and n_{y1} . (Actually, for reasons which will become clearer in the next section, the income effect in equation (3) should be ignored. We assume the commodity has a very small share in the consumers' total budget.) For example, if $q_2=0.5$, $h_3=.8$, and we ignore the income effect term, the Verdoorn values for n_m and E_{s1} imply that $s_{12}=4.0$ and $s_{a3}=.625$.

Is it reasonable to suppose that the elasticity of substitution is generally much greater between imports from different sources than between imports and domestic production? I would suppose that in many commodities, though certainly not all, the two elasticities of substitution would be about the same. This assumption will drastically alter our estimates of trade creation and trade diversion. But it is time to drop some of the highly restrictive assumptions of this section and develop a more general model. This is done in the next section.

II

In the model used in the present study, domestic production is divided into two parts—that which is competitive with imports (we call this M_3), and “all other commodities,” which we call M_o . We assume that the cross-elasticities of demand for M_1 , M_2 , and M_3 with respect to a

change in the price of M_o are all equal ($n_{10}=n_{20}=n_{30}$). The income elasticities are also equal ($n_{y1}=n_{y2}=n_{y3}$). These assumptions imply that the Slutsky cross-elasticities are equal ($n_{10}^*=n_{20}^*=n_{30}^*$) and that the utility function can be written

$$U[M_k(M_1, M_2, M_3), M_o]$$

where M_k is homogeneous.

Let us define:

h_i ($i=1, 2, 3$) = the share of M_i in the consumption of the commodity category in question

H_i ($i=0, 1, 2, 3$) = the share of M_i in total consumption

s_{k0}, s_{12}, s_{13} = elasticities of substitution.

Assuming for the moment that P_1 is the only price which changes, we can write the effects of a tariff preference on M_1 as follows (for more detail on the model, as well as on empirical procedures, see [3]):

$$(5) \quad M_1 = -P_1(h_2s_{12} + h_3s_{13}) - P_1h_1(H_0s_{k0}) - P_1H_1n_{y1}$$

$$(6) \quad M_2 = +P_1(h_1s_{12}) - P_1h_1(H_0s_{k0}) - P_1H_1n_{y2}$$

$$(7) \quad M_3 = +P_1(h_1s_{13}) - P_1h_1(H_0s_{k0}) - P_1H_1n_{y3}$$

In general, P_2 and P_3 will change. Appropriate terms must be added to the demand equations above to take account of these price changes.

In addition to the demand equations, we have three supply equations:

$$P_1 = \frac{\Delta t}{1+t} + \frac{M_1}{(1+t)e_1};$$

$$P_2 = \frac{M_2}{e_2}; \quad P_3 = \frac{M_3}{e_3}$$

where

$$s_{a3} = -\frac{\partial(M_o/M_3)^*}{\partial P_a} \frac{1}{M_o/M_3} = -\left[\frac{1}{M_a} \left(\frac{\partial M_a}{\partial P_a}\right)^* - \frac{1}{M_3} \left(\frac{\partial M_3}{\partial P_a}\right)^*\right]$$

Since $(\partial M_3/\partial P_a)^* = -(\partial M_o/\partial P_a)^*$, we can write

$$h_3s_{a3} = -\frac{M_3}{M_a + M_3} \left(\frac{\partial M_a}{\partial P_a}\right)^* \left(\frac{1}{M_a} + \frac{1}{M_3}\right) = -\left(\frac{\partial M_a}{\partial P_a}\right)^* \frac{1}{M_a} = n_m^*$$

If we hold P_a constant and let P_3 change, a similar derivation leads to the same result.

- e_i ($i=1, 2$) = the export supply elasticity of M_i
 e_3 = the domestic supply elasticity of M_3
 t = the tariff rate

We now have six equations, which can be solved for the six unknowns: \dot{M}_1 , \dot{M}_2 , \dot{M}_3 , \dot{P}_1 , \dot{P}_2 , \dot{P}_3 .

This formulation has the advantage of clarifying the appropriate size of the income effects. If the commodity in question occupies less than one percent of consumers' budgets, as is customary for import categories, the income effects will clearly be negligible, as equations (5), (6), and (7) show.

For lack of space, the empirical assumptions are treated here with utmost brevity (for more detail, see [3]). I will describe the procedure first for finished manufactures. I followed Balassa in accepting 4.12 as the U. S. import demand elasticity for finished manufactures. The appropriately weighted share of imports in domestic consumption is .201. I assumed that the domestic supply elasticity of finished manufactures was fairly high—I picked a value of 4.0. I also assumed a rather small degree of substitutability between the commodity in question and all other commodities— $H_o S_{ko}$ was taken at 0.5. These values imply an elasticity of substitution between imports and domestic production of 6.16. I used this same figure for the elasticity of substitution between preferred and non-preferred imports.

I assumed the elasticity of export supply from nonpreference-receiving countries is rather high—I used 6.0. I abstained from making any estimate of the export supply elasticity in the less developed countries themselves. The model calculates the outward shift in the demand curve facing less developed countries as a result of the granting of preferences. This outward shift in the demand curve would equal the additional exports only if the LDC export

supply curve were perfectly elastic. The additional exports corresponding to any particular elasticity of supply can of course easily be calculated.

For semimanufactures in the United States, the elasticity of substitution between imports and domestic production and between preferred and nonpreferred imports was taken to be the same as for finished manufactures—6.16. The elasticity of substitution between the commodity in question and all other commodities was assumed to be 0.2, and the domestic supply elasticity was taken as 0.59. (These values, combined with the appropriately weighted share of imports in domestic consumption of 0.27, imply an import demand elasticity for semifinished manufactures of 1.63.) The elasticity of export supply in nonpreference receiving countries was put at 1.5.

The parameters assumed for the other developed countries are listed in Table 1, with the U.S. values given for comparison. Since import-consumption ratios were not available for countries other than the United States, I followed Balassa in assuming these ratios to be .40 for the United Kingdom and .333 for the E.E.C. and Japan. They were assumed to be the same for all commodities.

I want to mention that I consider the assumptions for semifinished manufactures to be more arbitrary than those for fin-

TABLE 1—ASSUMED VALUES OF PARAMETERS

Finished Manufactures	U.S.	E.E.C., U.K., and Japan
s_{12}	6.16	4.0
Domestic supply elasticity	4.0	2.0
Competitors' export supply elasticity	6.0	6.0
Semi-Finished Manufactures		
s_{12}	6.16	4.0
Domestic supply elasticity	0.59	0.30
Competitors' export supply elasticity	1.5	1.5

TABLE 2—ESTIMATED EFFECTS OF FIFTY PERCENT PREFERENCES ON EXPORTS BY LESS-DEVELOPED COUNTRIES OF FINISHED AND SEMIFINISHED MANUFACTURES
(FIGURES IN MILLIONS OF U.S. DOLLARS)

	Importing Country	1965 Imports	Post-KR Imports*	Gain in LDC Exports		
				Total	Trade Creation	Trade Diversion
A. Finished Manufactures	U.S.	788.0	862.3	234.5	215.1	19.5
	U.K.	117.9	122.7	12.8	8.0	4.8
	E.E.C.	427.7	462.3	79.6	55.5	24.2
	Japan	64.9	70.3	5.6	4.2	1.3
Total		1398.6	1517.7	332.5	282.8	49.7
B. Semifinished Manufactures	U.S.	289.4	304.5	38.7	35.2	3.5
	U.K.	70.6	72.4	4.7	1.7	3.0
	E.E.C.	273.5	276.0	15.6	6.7	8.8
	Japan	66.2	68.2	4.0	2.2	1.8
Total		699.8	721.1	63.0	45.8	17.1
C. Total Manufactures	U.S.	1077.5	1166.8	272.9	250.2	23.0
	U.K.	188.5	195.1	17.5	9.7	7.8
	E.E.C.	701.2	738.3	95.3	62.2	33.0
	Japan	131.2	138.6	9.6	6.5	3.1
Total		2098.4	2238.7	395.5	328.6	66.9

* 1965 imports adjusted for the static effects of Kennedy Round tariff cuts. Detail may not add to total because of rounding.

ished. In general, we know little about supply elasticities. This deficiency in our knowledge is not too crucial in cases where the results are not very sensitive to the precise values of supply elasticities, and in finished manufactures, if we are correct in assuming supply elasticities to be generally high, their precise value does not matter much. But in semifinished manufactures, the estimates are quite sensitive to supply elasticities, in the ranges of the values that we have judged most likely.

Before giving our estimates of the effects of tariff preferences, we must describe the relevant trade flows. Of the \$25.6 billion of exports from less developed to developed countries in 1965, only \$7.38 billion were manufactures and semimanufactures, according to the UNCTAD definition. If we take out petroleum products, processed agriculture, and nonferrous metals, which were not included in the present study, we

are left with \$3.16 billion. Our trade data cover only \$2.096 billion out of this \$3.16 billion. The discrepancy is due to our omission of developed importing countries other than the U.S., the U.K., the E.E.C., and Japan, and our omission of certain products because either tariff or import data were not available on a four-digit SITC level.

The 1965 trade flows were adjusted to take account of the static effects of the Kennedy Round tariff cuts. The tariff preferences were assumed to be fifty percent of the post-Kennedy Round rates. The predicted trade gains in manufactures by less developed countries have not been adjusted for losses in exports of raw materials.

The estimates are shown in Table 2. The static effects of preferences on LDC exports are estimated at \$396 million, of which \$333 million are in finished manu-

factures. The United States accounts for \$273 million out of the total of \$396 million.

The estimates of trade diversion and creation are quite interesting. In finished manufactures, where supply elasticities are assumed to be high, trade creation exceeds trade diversion by a considerable margin. This occurs because domestic production of the particular commodity is usually large, relative to imports from developed countries, and because we assume that the elasticity of substitution between imports from developed and less developed countries is the same as that between imports from less developed countries and domestic production. In semifinished manufactures, trade diversion exceeds trade creation in some countries. As mentioned earlier, the estimates for semifinished manufactures are quite sensitive to supply elasticities, and our assumption that the export supply elasticity of competitors is considerably larger than the domestic supply elasticity accounts in part for these results.

III

The main points of this paper can be briefly summarized. The Verdoorn-Johnson model of customs unions can be interpreted in terms of consumer behavior under a separable utility function. Such an interpretation places restrictions on the acceptable values of the Verdoorn-Johnson elasticities—the share elasticity of substitution and the elasticity of import demand. In particular, the values of 2.0 and 0.5 for these two elasticities are too far apart. Bringing them closer together increases the estimated trade creation at the expense of trade diversion.

This paper also presented a somewhat more general model of tariff discrimination than the Verdoorn-Johnson one. In our model, if supply elasticities are generally high, if the elasticity of substitution between partner and nonpartner imports is the same as that between partner imports

and domestic goods, and if import-consumption ratios are small, trade creation will be considerably larger than trade diversion. Our model thus provides a possible explanation for the apparent fact that trade diversion has been rather small relative to trade creation in the European Common Market.

REFERENCES

1. Lloyd Atkinson and John E. Morton, "Estimating Nonmember Losses in the Formation of a Customs Union," Univ. of Michigan, March 14, 1968, unpublished paper.
2. Bela Balassa, "Trade Creation and Diversion in the European Common Market," *Econ. J.*, March 1967, 77, 1-21.
3. Christopher K. Clague, "The Trade Effects of Tariff Preferences for Less-Developed Countries," Mimeo, Jan. 1971.
4. F. J. Clavaux, "The Import Elasticity as a Yardstick for Measuring Trade Creation," *Economia Internazionale*, Nov. 1969, 22, 606-612.
5. S. M. Goldman and H. Uzawa, "A Note on Separability in Demand Analysis," *Econometrica*, July 1964, 32, 387-398.
6. L. H. Janssen, *Free Trade, Protection and Customs Union*, Leiden, 1961.
7. Harry G. Johnson, "The International Competitive Position of the United States and the Balance of Payments Prospect for 1968," *Rev. Econ. Stat.*, Feb. 1964, 46, 14-32.
8. Lawrence B. Krause, *European Economic Integration and the United States*, Washington, 1968.
9. Mordechai E. Kreinin, "Trade Creation and Diversion by the EEC and EFTA," *Economia Internazionale*, May 1969, 22, 273-280.
10. Edwin M. Truman, "The European Economic Community: Trade Creation and Trade Diversion," *Yale Econ. Essays*, Spring 1969, 9, 201-257.
11. P. J. Verdoorn, "The Intra-Bloc Trade of Benelux," pp. 291-329 in E.A.G. Robinson, ed. *The Economic Consequences of the Size of Nations*, New York, 1960.

Nontariff Barriers and the Export Performance of Developing Economies*

By INGO WALTER
New York University

The problem of expanding, diversifying and stabilizing the export earnings of less-developed countries has been of major interest both to development planners and to those concerned with the evolution of international trade relations. A central determinant of LDC export performance and growth is their ability to secure expanded and reliable access to the major developed-country markets for manufactures and semimanufactures—access that is governed in large measure by the trade policies pursued by these countries and successes achieved in their liberalization.

It is convenient to group the instruments of trade policy into tariff and nontariff measures, respectively, and to consider as the primary focus for market access the developed countries of Western Europe, North America and Australasia.¹ Tariff levels maintained by these countries have in general declined markedly over the last two decades, and are presently subject to further alteration under which they accord the developing countries a measure of vertical equity in their trade policies in the form of the UNCTAD generalized tariff preferences. Nevertheless, the developing nations frequently maintain that much of the recent

tariff liberalization—and even the national and group preference offers—have tended to exclude products of major actual or potential export interest to them, and that effective rates of protection have remained high due to wide differentials in rates of duty as applied to imports of products at different levels of processing.

At the same time, measures other than tariffs impeding access to international markets have come in for mounting scrutiny, both as potentially viable alternative sources of protection for import-competing interests and as increasingly important targets for those advocating continued trade liberalization. There is little question about the *relative* increase in importance of nontariff obstacles to trade in recent years. Allegations charging significant proliferation of such obstacles—as the efforts of protection-seeking forces are redirected from the relatively rigid and internationally bound national tariff structures to a rather more pliable and multifaceted complex of nontariff barriers (NTB's)—have yet to be substantiated. The present paper focuses on the implications of nontariff measures, applied to imports by the industrial countries and bearing on manufactures and semimanufactures, for the export performance and prospects of the developing economies.

I. Nontariff Barriers

The character and operation of nontariff obstacles to international trade has been considered in some detail.² Such

* This paper was written while the author was a Senior Fellow at the Center for International Studies, New York University. Helpful comments by Robert G. Hawkins on earlier drafts are gratefully acknowledged.

¹ For purposes of this paper, we define the developed market economy countries to include the following: Canada, the United States, Australia, New Zealand, Japan, the EEC member countries, the United Kingdom, Finland, Norway, Sweden, Denmark, Austria, and Switzerland.

² See Robert E. Baldwin, *Nontariff Distortions of International Trade* (Washington, D.C.: The Brook-

measures may be defined to include all governmental policies and practices which serve to distort the volume, direction or product-composition of international trade. For certain purposes it is useful to differentiate between those types of nontariff measures, (a) which are generally employed as instruments of commercial policy, e.g., quotas, subsidies to import-competitors, and variable levies; (b) which are sporadically employed with trade-restrictive intent, e.g., packaging and marking requirements, phytosanitary regulations, customs valuation and classification practices; and (c) whose trade-restrictive effects are largely incidental to nontrade-policy objectives, e.g., certain types of consumption taxes and government monopolies.³ In addition, for analytical reasons it is advisable to group NTB's into those which can be clearly identified as bearing on specific products (e.g., quantitative import controls), and those bearing on all or broad groups of products (e.g., government procurement).

Nontariff obstacles affecting *imports* tend to influence the volume and pattern of trade by quantitative limitations or by directly affecting relative costs and prices. In either case, their restrictive impact may be felt at the point of importation or at several other points in the relevant credit-distribution-marketing chain. Un-

certainities created by such measures—attributable for example to a lack of information—bear additionally on importers and distributors as well as on foreign exporters and suppliers, and tend to reinforce the primary effects of the nontariff obstacles themselves.

A much more limited set of measures operates on the *export* side, again functioning either by quantitative means or by influencing relative prices and costs, with the intent either of artificially stimulating the volume of exports through implicit or explicit subsidization or of restricting foreign sales of specific commodities.

Evaluation of the trade-distortive impact of nontariff measures—either in the aggregate or with respect to specific products—has proved to be particularly difficult. Several nontariff obstacles may affect a single product or group of products, with some NTB's affecting all or entire categories of imports. Together with the applicable tariff rates, such *multiple-stacking* of obstacles makes it virtually impossible to assign responsibility for trade-restrictive impact or to determine which measures may be redundant. Certain NTB's are highly variable over time or depend entirely upon discretionary action on the part of the authorities concerned, while others are applied in a clandestine manner. In most cases the availability of data on their specific administration is highly restricted.

The *a priori* impact of most types of NTB's bearing on imports is nevertheless relatively clear. Fixed cost-imposing NTB's operate in much the same way as do tariffs, except that the uncertainty factor may play a major role and in most cases there are marked differences in the revenue effects. Import charges, which are systematically or randomly variable pose substantially greater difficulties for analysis, their effects being in some respects similar to those of quotas and in others

ings Institution, 1970) or the author's "Nontariff Barriers and the Free Trade Area Option," *Banca Nazionale del Lavoro Quarterly Review*, March 1969.

³ The measures falling into the first group tend to lend themselves most readily to bargaining-type solutions, while those included in the second group tend to be susceptible to international standardization as a liberalization technique. However, as demonstrated by the agreement to eliminate piston-displacement taxation of motor vehicles by certain European countries in return for the elimination of the American Selling Price system of customs valuation negotiated during the Kennedy Round, this need not always be the case. The trade-distortive effects of measures falling into group (c) are considerably more difficult to eliminate due to their basically non-trade-related goals.

more akin to tariffs while generally incorporating a major element of uncertainty. The impact of selective subsidies to import-competing suppliers is likewise similar to that of tariffs as regards the volume of trade and the national economic structure, but quite different with respect to the fiscal aspects, domestic prices and implications for effective protection and demand levels.⁴ While the economics of global import quotas have been rather thoroughly considered, the impact of other types of quantitative restraints such as buyers' quotas, domestic procurement practices, advertising restrictions and other marketing controls are less well explored. Many of these obstacles lend themselves to analysis only on a highly detailed, case by case level and defy generalization.

Finally, it should be noted that the application and administration of nontariff measures in the industrial countries is generally highly complex. Certain types of measures tend to be primarily legislative matters, while others hinge on the policies of numerous executive agencies, often with little apparent central coordination or systematic regulation. This aspect tends to facilitate the availability of protection via NTB's to import-competing interests. It also removes the issue at least partly from public debate and scrutiny.

⁴ The question of similarities and differences in the effects of tariffs, quotas and subsidies has been explored in the literature. See for example W. M. Corden, "Tariffs, Subsidies and the Terms of Trade," *Economica*, August 1957, and Michael Michaely, "A Note on Tariffs and Subsidies," *American Economic Review*, September 1967. On quotas, see J. Bhagwati, "On the Equivalence of Tariffs and Quotas," in R. E. Baldwin *et al.*, *Trade, Growth, and the Balance of Payments*, (Chicago: Rand McNally, 1965) and his "A Note on the Equivalence of Tariffs and Quotas," *American Economic Review*, March 1968; H. Shibata, "A Note on the Equivalence of Tariffs and Quotas," *American Economic Review*, March 1968; M. E. Kreinin, "More on the Equivalence of Tariffs and Quotas," *Kyklos*, Fasc. I (1970) and the author's "Comment" in *Kyklos*, Fasc. I (1971).

II. *Impact On Developing Countries*

In addition to the standard arguments concerning the impact of NTB's, the case is often stated that, even when nominally nondiscriminatory, such measures may in fact impose an extraordinary burden on the developing countries, either individually or as a group, and that special consideration should therefore be accorded their interests in liberalization efforts.⁵ The alleged *de facto* discrimination is attributed to two basic sources.

First, severity of the impact on the developing countries of a variety of nontariff obstacles to imports tends to be disproportionate in comparison with their impact on competing developed countries. The basis for this argument includes the following: (a) the affected suppliers often are relatively less able operationally to cope with such obstacles than their competitors in advanced countries, while their capability of affecting the requisite adjustments is notably more limited; (b) more generally, alternative production possibilities tend to be restricted, or essentially nonexistent, for the developing economy; (c) from an informational standpoint, developing countries often do not have sufficient resources to determine precisely the nature of existing nontariff measures and the mechanics of their administration, resulting in faulty and costly supply responses, a high degree of uncertainty, and a corresponding dampening of the export effort; (d) in certain product areas extraordinarily rigorous standards and procedures may be applied to imports from developing countries, particularly with respect to health and quality controls implemented as part of a concerted national policy or at the discretion of the officials involved;

⁵ For instance, see the debates in the fourth (1970) session of the UNCTAD Committee on Manufactures, as noted in the proceedings (UNCTAD documents TD/295-TD/B/C.2/97).

and (e) as in the case of past tariff liberalization efforts, the interests of developing countries have tended to assume a secondary role in ongoing intergovernmental attempts to deal with nontariff obstacles.

With respect to the last point, it may be argued that particularly the GATT procedure—relying on notifications by member countries on nontariff measures affecting their exports—places the developing countries at a fundamental disadvantage. This notion, in addition to their inherently weak bargaining position and the fact that many developing countries are not contracting parties to the General Agreement, underlies the position that nontariff obstacles of especial concern to the developing countries may not receive adequate emphasis in this particular international forum.⁶

Second, besides the possible disproportionate impact of individual nontariff obstacles as between LDC's and the advanced countries, it may be that identifiable nontariff obstacles more often fall on products or product-groups of particular export interest to the developing countries, given their supply capabilities, than they do on manufactured and semimanufactured products whose supply falls largely in the province of the industrial countries. That is, manufactured and semimanufactured products most frequently subject to nontariff barriers tend to be precisely those in which the developing countries presently have, or are likely to establish in the near term, a competitive advantage in the international market.

In the case of both sets of arguments, although they are intuitively appealing and basically supportive of the developing countries' position, rigorous and empirically defensible verification is elusive at

best. However, a number of lines of inquiry may usefully be pursued to shed some additional light on the matter.

III. *Disproportionate Impact*

With respect to the first line of reasoning—that limited adjustment capabilities, greater vulnerability to uncertainty, and more restricted knowledge of the nature and application of NTB's generally result in a disproportionate impact on developing-country suppliers compared with their developed-country competitors—one must continue to rely almost entirely on qualitative and subjective evaluation. Yet this aspect is important. If true, the developing countries would be justified in seeking *preferential* relief from specific NTB's on grounds at least as persuasive as the standard arguments for trade preferences. Of the 38 more or less distinct types of NTB's identified,⁷ at least 17 would appear capable of exerting a marginally heavier impact on developing-country suppliers.

The findings of a series of six detailed product studies seem to suggest the following:⁸ In the case of quantitative import restrictions, any such differential

⁷ For the relevant NTB breakdown, see *Liberalization of Tariffs and Nontariff Barriers*, Document TD/B/C.2/R.1, (Geneva: UNSCAD, 10 December 1969), pp. 10–12.

⁸ The following manufactured and semimanufactured products were considered:

- a) prepared and preserved meats (SITC 012.9/013.8/013.3);
- b) prepared and preserved vegetables (SITC, 055.1/055.51/055.52);
- c) household ceramic articles (SITC 666.4/666.5/666.6);
- d) preserved and prepared fruits and juices (SITC 053.3/053.5/053.9);
- e) starch derivatives and products (SITC 055.44/599.51/055.04); and
- f) sugar confectionary and chocolate (SITC 062.01/073.0).

Cf., R. L. Allen and I. Walter, "An Analysis of the Impact of Nontariff Measures Imposed by Developed Market Economy Countries on Representative Products of Export Interest to Developing Countries," UNCTAD Secretariat Working Paper (Mimeo., 1970).

⁶ Among the developing countries that are Contracting Parties to the GATT, only a few such as Brazil, Mexico, Yugoslavia, India and Pakistan participated meaningfully in the initial GATT exercises on nontariff barriers.

may be attributed to the administration of quotas and discretionary licensing of imports. The issuance of import permits is frequently undertaken primarily in favor of "traditional" sources of supply and importers having long-standing relationships with certain foreign producers, limiting market access by newly emerging suppliers in the developing countries. Permissible quota sizes may be unknown or subject to activation under a variety of explicit or implicit trigger mechanisms. The quantity of allowable imports be highly variable over time, import-calendars may be employed, or "buyers' quotas" may be used under which each allocation is too small to service efficiently. Such administrative techniques frequently appear to result in unfilled quotas as developing-country suppliers, faced with a high degree of uncertainty, are frustrated in their efforts to gain improved and stable market access.

Even relatively liberal or quasi-automatic licensing of imports on occasion seems to impose a significant burden on developing countries due to complexities involved in securing import permits, attendant costs, delays in issuance, variations in the period of validity, and disclosure requirements as to final purchaser and price. "Voluntary" export restraints, backed by the threat of mandatory import controls, as yet seem to affect relatively few developing countries or product groups of export interest to them. Current application of this device to various textiles, netting, plastics, porcelain articles, cutlery, and electronic components may portend future intensification, and the limited bargaining power of the individual developing countries would leave them little choice but to acquiesce.

In the case of public and quasi-public procurement, limited publication of tenders, short bidding times, wide discretion allowed purchasing authorities, and de-

velopment aid tied to purchases in the donor country appear to have a disproportionate effect, but in a relatively narrow range of products. At the same time, export restrictions intended to deny raw materials or intermediates to foreign processing industries—in order to protect domestic producers of the finished product—seem to have had a deleterious effect on certain developing countries with regard to textiles, and may be subject to future intensification in connection with offshore manufacturing in low-labor-cost areas by firms based in the industrial countries.

There are a number of other NTB's whose impact in individual cases also appears to be asymmetrical. Variable import levies and surcharges imposed by the EEC and a number of individual developed countries on processed agricultural products frequently involve abrupt shifts in magnitude and considerable difficulties in calculating the precise amount of the levy, rendering almost impossible advance determination of import charges. Advance deposit requirements and credit restrictions on purchases from abroad may fall more heavily on importers with weak financial backing, while developing countries are in a fundamentally disadvantageous position in competing with direct or indirect subsidies granted to exporters or import-competitors and their suppliers in developed nations.

A similar pattern is found in the case of nontariff obstacles not necessarily implemented with protective intent. Technical standards imposed on a variety of products (e.g., plywood, fertilizers, electronic components) appear to be a problem, although packaging and labelling regulations, safety standards and phyto-sanitary requirements are frequently much more striking in this regard. For developing-country suppliers, variations in standards and the requisite short packaging runs and high tooling and design costs can rep-

resent a heavy burden—especially when exclusivity in packaging or labelling is enforced—a problem that is often compounded by a lack of precise information concerning the specific nature of such regulations. Health requirements are of even more serious concern for developing countries, specifically with regard to processed food products. It is difficult for them to cope with variations in national regulations concerning permissible content of insecticides, certain chemicals, food additives, coloring and bacteria count, problems compounded by nonrecognition of national inspection standards and occasional explicit discrimination at the customs frontier.

Finally, the standard argument concerning indirect taxation and border tax adjustments may also carry unusual implications for developing countries, although the extent of any such asymmetry remains subject to doubt. Not only do most developing countries fail to apply a system of drawbacks of charges bearing on exported products—thereby failing to meet one of the conditions for nondistortive application of the “destination principle” of taxation in international trade—but the fiscal exigencies of individual LDC’s in addition often dictate the imposition of taxes on exports. It can therefore be argued that manufactured and semi-manufactured products imported from certain developing countries bear an extraordinarily heavy tax burden on final sale.

More precise examination of the impact of nontariff measures on exports of developing countries must of necessity focus on in-depth studies of narrowly-defined product groups, incorporating as many factors as possible operating on the supply and demand side, in an attempt to isolate their trade-distortive implications in an effective-protection framework. In a very preliminary manner it was estimated that

in the absence of NTB’s bearing on the above-mentioned six illustrative product groups, imports by the developed market-economy countries from developing areas might have been \$750–\$820 million as opposed to an actual 1968 trade base of \$486.4 million. A number of methodological difficulties and data constraints were encountered which necessarily limit the validity of the results, and some rather drastic assumptions about the supply capabilities of the LDC’s had to be made; the calculated 54–68 percent margin should therefore be taken only as an indication of the order of magnitude that may be involved in estimates of the restrictiveness of NTB’s, particularly in such important and highly sensitive sectors as processed agricultural products. Nor does it follow that such estimates necessarily indicate the possible magnitude of trade expansion that might occur if NTB’s applied to these products were liberalized. In some instances such supply-related factors as domestic absorption, quality control, marketing techniques, and the developing countries’ own import restrictions were judged even more serious impediments to increased exports than the nontariff or tariff obstacles facing them in the developed countries.

IV. *Differential Incidence*

The second line of reasoning—that there exists a great deal of overlap as between manufacturers and semimanufacturers particularly subject to NTB’s and products of export interest to developing countries—lends itself somewhat more readily to investigation. One approach at a rather aggregative level involves cataloging the known nontariff measures which can be identified as bearing on manufactures and semimanufactures⁹ ap-

⁹ As defined in UNCTAD document TD/B/C.2/3 of 2 July 1965 to include all SITC product groups in sections 5–8 as well as a variety of processed foods,

plied by each of the developed market-economy countries and the EEC,¹⁰ classified in a manner coincident with available (SITC) trade data.

In 1968, the developed market economy countries imported approximately \$128 billion in manufactures and semimanufactures, of which about 28 percent (\$36 billion) were subject to known NTB's. In the same year, 33 percent of the \$21 billion of manufactured and semimanufactured imports from developing countries entered under such obstacles. Put another way, the developing countries' share of manufactured and semimanufactured imports of the developed countries in 1968 was 16.5 percent, but their share of imports subject to NTB's was 20.9 percent. Acceptance of the earlier proposition that certain NTB's on individual products or product-groups bear disproportionately on developing-country suppliers and hence tend to bias downward their actual market share would mean that the *hypothetical* latter figure is even larger relative to the former. But even assuming the imposition of nontariff measures had no effect on relative import shares, it would appear that product groups for which developing countries hold a relatively large share of developed-country imports are also characterized by a comparatively intensive application of NTB's—and hence that imports from developing countries do indeed tend to be somewhat more heavily subject to such measures than do developed-country imports in general. Table 1 indicates, however, that there are substantial variations in the figures as among the individ-

TABLE 1—1968 IMPORTS FROM DEVELOPING COUNTRIES
SUBJECT TO NTB'S: MANUFACTURES
AND SEMIMANUFACTURES

Imports of:	Share of Developing Countries in Total Imports (percent)	Share of Developing Countries in Imports Subject to NTB's (percent)
U.S.	18.6	23.5
Canada	6.8	3.5
Australia	12.5	7.9
New Zealand	10.8	2.0
Japan	43.2	58.4
U.K.	22.9	13.9
Norway	7.5	2.2
Sweden	9.1	6.2
Denmark	7.2	4.7
Finland	2.7	6.8
Austria	3.0	3.5
Switzerland	5.5	13.1
EEC (from non-member)	7.5	13.7
Belgium-Luxembourg	15.0	27.6
France	17.9	37.9
F.R. Germany	17.0	6.9
Italy	25.0	4.3
Netherlands	11.7	4.3
Total	16.5	20.9

Data: Classification of non-tariff barriers by application of products grouped by SITC 5-digit identification, see *Liberalization of Tariff and Non-Tariff Barriers*, loc. cit., Annex II. Trade data: U.N. Statistical Papers, Series D, *Commodity Trade Statistics* (New York: U.N., 1969); and OECD Statistical Papers, Series C, *Foreign Trade Statistics* (Paris, OECD, 1969). Definition of manufactures and semi-manufactures, see UNCTAD document TD/B/C.2/3 of 2 July 1965.

ual markets, and what is true for the total holds for only half of the individual importing areas.

There is also a good deal of variation in the incidence of NTB's applied by the advanced countries among products.¹¹ This is indicated by an examination of the application of nontariff measures to imports of major manufactured and semimanufactured product groups for each of 17 developed countries under consideration,

¹¹ For various methods of measuring the incidence of NTB's, see source cited in note 7, pp. 42-46. The most serious limitation of all such measures is that they do not differentiate between the various types of nontariff obstacles according to their restrictiveness of trade.

beverages, fuel and timber products in sections 0-4. See U.N. Secretariat, Statistical Papers, Series M, No. 34, *Standard International Trade Classification, Revised* (New York: U.N., 1961).

¹⁰ Nontariff measures which affect all product groups and hence cannot be identified as bearing on specific items are ignored here, although some of these measures may affect certain product groups more than others.

using as a measure of incidence the proportion of individual items under a given product heading subject to NTB's, weighted by OECD imports.

Data for some representative countries are given in Table 2. For the U.S., Japan, and France, for example, the incidence of NTB's on imports from developing countries is significantly higher than on imports in general. In the U.S. case, this is due to concentration of NTB's in petroleum products and textile garments; in Japan it may be attributed to heavy applications of such measures on these same groups plus beverages and tobacco, certain chemicals and textile yarns and fabrics, while the French case is mainly attributable to a heavy NTB incidence on processed foods, beverages, tobacco and wood products. Even in countries where the overall incidence of nontariff measures on developing countries is not significantly higher or is even markedly lower than average (e.g., F.R. Germany), instances of differential incidence always exist in individual product groups, particularly various light manufactured articles, semifinished products and intermediates, processed agricultural products, beverages and tobacco.¹²

It is also interesting to note that the intensity of application of NTB's appears to correspond at least ordinarily to the degree of competitive advantage the developing countries are deemed to possess in the production of manufactures and semi-manufactures for the international market. Table 3 ranks products by calculated LDC competitive position and indicates the extent of NTB application by all de-

veloped market economy countries combined.¹³ The 60 product groups represent categories where some comparative advantage is thought to exist; of the 63 remaining manufactured product groups at this level of aggregation to which no foreseeable LDC advantage is ascribed, only 15 evidenced susceptibility to NTB's. When the NTB factors are also ranked, the correlation (Spearman) coefficient between the two rankings is 0.2317, which is statistically significant at the .95 level of confidence. Hence one concludes that those manufactured and semimanufactured product groups for which a relatively strong LDC competitive position exists for the developing countries also tend to be the ones most heavily subject to nontariff applications—assuming adequate reliability in the measurement techniques employed.

V. Conclusions

We have tried to indicate in this paper that there is considerable preliminary evidence to suggest that the developing countries are rather susceptible to nontariff measures applied to imports by advanced countries and that, indeed, the structure of protection and quasiprotection indicated by these obstacles may well be systematically—if unintentionally—biased against them. This would support similar conclusions by other writers¹⁴ concerning

¹² The competitiveness factors are taken from a recent study by H. F. Lydall for UNCTAD, *The Performance of Developing Countries as Exporters of Manufactures to the Developed Market Economy Countries*, Document TD/B/C.2/91 (Geneva: UNCTAD, 22 December 1969). Because of high reliance on natural resource endowments, petroleum products and nonferrous metals are excluded in the competitiveness estimates, which are based on human capital-intensity, physical capital-intensity and the size of plant as determining variables. The NTB factors represent an enumeration of the specific NTB's bearing on individual product groups in the developed countries. See *Liberalization of Tariffs and Non-Tariff Barriers*, loc. cit., Annex II.

¹⁴ See Bela Balassa, *The Structure of Protection in*

¹³ It should be noted that tariff rates and NTB's do not appear systematically to complement or substitute for one another in national protective structures: preliminary analysis of post-Kennedy Round tariff averages and NTB incidence factors at the SITC 2-digit level yielded significant positive or negative correlation in the case of few of the developed countries or customs areas under consideration here

TABLE 2—INCIDENCE OF NON-TARIFF MEASURES ON IMPORTS OF MANUFACTURES AND SEMIMANUFACTURES, SELECTED COUNTRIES, 1968*

SITC No.	Product Description	United States		United Kingdom		Japan		EEC External†		F.R. Germany‡		France‡	
		Total	LDC	Total	LDC	Total	LDC	Total	LDC	Total	LDC	Total	LDC
Ex. 0	Processed foods	21.4	9.2	27.2	28.6	100.0	100.0	90.2	77.7	62.8	50.5	70.7	81.1
Ex. 1	Beverages and tobacco	93.3	92.5	56.9	32.8	92.7	98.7	93.0	98.9	71.6	97.3	91.4	98.7
Ex. 2	Crude materials	0.2	—	—	—	100.0	100.0	—	—	—	—	44.5	83.8
Ex. 3	Fuels and lubricants	97.6	99.5	—	—	70.9	85.4	—	—	—	—	100.0	100.0
Ex. 4	Oils and fats	32.0	—	—	—	—	—	76.1	0.6	—	—	44.2	—
Ex. 5	Chemicals	69.4	31.6	1.9	1.4	55.1	83.9	16.0	17.5	5.1	6.8	33.7	25.5
6	Manufactures classified chiefly by material	23.5	6.4	31.0	15.3	35.6	86.6	—	—	16.8	14.9	19.6	11.8
7	Machinery and transport equipment	16.5	1.4	17.8	34.0	45.0	36.8	—	—	28.1	44.7	39.1	61.6
8	Misc. Manufactures	40.5	64.7	29.6	54.3	17.8	7.4	—	—	23.8	52.0	47.9	68.3
9	Products not elsewhere classified	78.5	15.4	100.0	100.0	—	—	—	—	3.4	—	—	—
Ex. 0-9	Total manufactures and semimanufactures	44.3	66.9	19.3	9.6	35.6	57.7	3.0	3.9	18.0	10.6	43.8	72.7

Data: See Table 1.

* Percent of manufactured and semimanufactured products in the indicated product groups, weighted by OECD imports, covered by each area's non-tariff measures, total and from developing countries in 1968.

† Non-tariff measures bearing on imports from nonmember countries only.

‡ Excluding EEC non-tariff measures bearing on imports from nonmembers.

TABLE 3—MANUFACTURES AND SEMIMANUFACTURES OF EXPORT INTEREST TO DEVELOPING COUNTRIES: COMPETITIVENESS AND NONTARIFF BARRIERS

SITC No.	Product Description	Competitive Position ^a	NTB Factor	SITC No.	Product Description	Competitive Position ^a	NTB Factor
841	Clothing	19.0	71.1	013	Preserved meats	5.2	78.0
657	Carpets	18.0	25.0	012	Dried and salted meats	5.2	97.5
243	Shaped wood	16.3	15.0	724	Telecommunications apparatus	4.9	42.0
899	Other products ¹	16.0	8.3	599	Other chemicals	4.7	10.5
831	Travel goods	13.9	0.0	431	Oils, fats and waxes	4.7	7.1
653	Woven noncotton fabrics ²	13.2	60.3	696	Cutlery	4.5	2.1
032	Preserved fish	12.8	37.5	665	Glassware	4.1	1.7
897	Jewelry	12.5	25.0	551	Essential oils	4.0	5.0
632	Wood products	12.5	2.7	821	Furniture	3.5	0.0
651	Yarn and thread	12.4	19.4	521	Tar, etc. from coal	3.4	6.7
656	Bags, sacks, linens	12.2	24.3	072.3	Cocoa butter	3.2	15.0
053	Preserved fruit	12.2	47.5	073	Chocolate	3.2	34.7
055	Preserved vegetables	12.2	46.3	071.3	Coffee extracts	3.0	44.6
894	Toys and sporting goods	11.8	6.7	661	Lime, cement, etc.	2.8	0.0
532	Tanning and dying extracts	11.5	2.5	541	Medicinal products	2.8	55.5
655	Special textile fabrics ³	10.0	15.6	861	Scientific instruments	2.7	4.8
663	Mineral manufactures nes	9.3	0.1	641	Paper and paperboard	2.3	4.7
851	Footwear	9.2	17.5	513	Inorganic chemicals, oxides	2.3	10.8
652	Woven cotton	7.8	117.1	514	Other inorganic chemicals	2.3	10.5
099	Other food products	7.6	21.3	629	Rubber products	2.2	10.2
052	Dried fruit	7.5	23.5	251	Pulp and waste paper	2.0	0.0
892	Printed matter	7.5	5.0	266	Synthetic fibers	2.0	4.3
893	Plastic products	7.5	5.0	678	Iron, steel tubes & pipes	1.7	23.0
631	Veneers and plywood	6.6	0.8	512	Organic chemicals	1.6	18.2
611	Leather	6.6	3.3	671	Pig iron	1.4	0.0
697	Household metal products	6.5	1.9	673	Iron and steel bars	0.1	30.0
698	Other metal products	6.4	1.5	674	Iron & steel plates, sheets	0.1	30.0
561	Manufactured fertilizers	5.9	23.0	533	Pigments, paints, varnishes ⁴	<0.1	4.3
812	Light and sanitary fittings	5.7	8.8	642	Art. of paper, pulp, paperboard ⁴	<0.1	0.0
662	Clay products	5.7	4.4	712	Agric. machinery, implements ⁴	<0.1	0.7

Data: Estimates of competitive position: UNCTAD, *The Performance of Developing Countries as Exporters of Manufactures to the Developed Market Economy Countries*, Doc. TD/B/C.2/91, 22 December 1969. Estimates of the NTB factor calculated from data assembled by the author and published in UNCTAD, *Liberalization of Tariff and Nontariff Barriers*, Doc. TD/B/C.2/R.1, 10 December 1969, Annex II, averaged (unweighted) over products included under indicated SITC category.

¹ Mainly basketwork, brooms, umbrellas, buttons, artificial flowers and wigs of human hair.

² Primarily jute fabrics.

³ Mainly cordage, rope and twine.

⁴ Categories where a longer-range competitive advantage may exist but not yet evident in the calculations.

⁵ There are 63 additional categories of products at this level of aggregation for which developing countries are not deemed to possess a comparative advantage under existing conditions and hence do not appear in this listing; of these, 15 are subject to NTB's and the remainder (48) are not.

tariff levels and structures, and derives from (a) the apparent coincidence of

the Industrial Countries and its Effects on the Exports of Processed Goods from Developing Countries UNCTAD Document TD/B/C.2/36, 25 May 1967; and Hal B. Lary, *Imports of Manufactures from Developing Countries* (New York: National Bureau of Economic Research, 1968).

manufactured and semimanufactured product groups subject to such obstacles and corresponding product groups of export interest to the developing countries; and (b) the disproportionate impact of certain NTB's on emerging suppliers located in the developing countries. With the implementation of the generalized

preferences, such considerations will tend to become rather more significant, a factor further accentuated by the essentially "nontariff" escape mechanisms built into a number of national or group preference offers.

There are, of course, broad questions concerning the implications of such conclusions for development. Is the orderly, sequential promotion of economic growth in the LDC's in fact seriously impeded by

such measures, or does their impact fall primarily on static allocational efficiency? Do the present and prospective "rules of the game" as implied by the structure of tariff and nontariff protection materially alter development strategy in the direction of greater inward orientation? Such aspects need to be considered in detail before a reliable judgement can be rendered with regard to the actual meaning of NTB's for the developing countries.

DISCUSSION

HARALD B. MALMGREN: The papers before us cover questions in development policy which normally do not receive adequate attention. There has been a concentration of both economic analysis and international negotiating effort, almost to the point of obsession, on the scheme of general tariff preferences formulated in the OECD and UNCTAD over the last few years. Preferences will be useful, but their effects will not be very large for yet some time. Professor Clague's estimates suggest only modest trade improvement. In the meantime, almost nothing has been done to analyze the wide range of nontariff barriers (NTB's) which are of much greater significance to the developing nations.

Professor Walter's interesting and constructive paper concentrates on NTB's maintained by developed countries on imports of manufactures. He observes that these NTB's fall relatively more severely on imports of developing nations. This is not surprising. If anything, I would have thought the discrimination would be even greater. He is correct in arguing that any NTB is more difficult for a developing country exporter to cope with than an exporter in a rich country who has the resources, marketing apparatus, and quality control experience to get around the impediments, or else jump them by building part of the production apparatus in the market abroad. I do not on this account agree, however, that NTB's should be liberalized *preferentially*. Many of them derive from varying national standards, taxation, labeling requirements, customs and administrative procedures, and so on. Most NTB's are tied to domestic social objectives, policies, and laws. For many of these, the basic problem is to bring some harmony into national practices through better international rules and consultative procedures. The rules should aim at preventing the passing to foreign interests of the costs of domestic policies, or where such cost-passing occurs, providing for compensation or other balancing adjustments. The most pernicious restrictions are the import quotas (often of the so-called "voluntary export restraint" type). These usually discriminate against developing

nations for two reasons: (1) the developed nations avoid putting quotas on each other's trade since each has the power of real retaliation, and (2) the politically sensitive import problems usually arise in the case of labor-intensive products from low-wage countries. A new framework for liberalizing quotas multilaterally is sorely needed, and a framework or guidelines to control *potential* restrictions is even more badly needed. Particularly useful would be new concepts or definitions of what constitutes "market disruption" and excessive labor displacement by imports, so that practical guidelines might be negotiated which would limit the politically capricious behavior of the rich countries regarding so-called "sensitive" imports.

On this business of rationalizing NTB's, the same need exists in the case of the developing nations. They maintain an amazing array of trade barriers. Most of these harassment devices are not measurable. If all these impediments could be simplified and converted to tariffs, there would be several benefits. The impediments could be measured and manipulated as instruments of development policy. The use of tariffs would provide revenue for governments, in the place of high administrative costs now borne in the administration of complex procedures, and in place of the bribery which is a natural partner of complexity. On this Professor Balassa has nothing to say, and perhaps in the future he might broaden his analysis to cover these matters.

Professor Walter's paper has a conspicuous omission, in leaving out agricultural NTB's such as the European Common Agricultural Policy, the Japanese rice program, and the tropical products taxes in Europe. These agricultural problems are becoming increasingly painful for all trading nations, with the poorer countries least able to cope with the competition of subsidies, and of protectionism.

There is increasing global recognition that the headlong plunge into import substitution in many countries has been unwise. Professor Balassa suggests some sensible guidelines for the trade policies of developing nations, both for those starting out on industrialization and

for those already well along the path of import substitution. His suggestion of dealing with the employment problem directly, through specific incentives for labor-using techniques, rather than through trade protection, is worthy of much attention, especially taking into account the rising unemployment problem in many developing nations. The temptation will be even greater in the next decade or two to use import protection to preserve employment opportunities in the developing nations, just as that temptation is growing in the U.S. and other rich countries. I also agree with his presumption that promotion of manufacturing should on balance be favored. His suggestion of equalizing the rates of protection, incentives for home use, and export subsidization, for all industries past the infant-industry stage is sound enough. Professor John Power has also been writing along these lines recently. However, just when an industry grows out of the infant stage is a real question. The definition of the phases of the process of maturity into an adult industry is itself a complex theoretical task. How much protection, at what stages? The declining scale of protection he suggests may be the only answer, but it will be hard to sell.

Professor Clague's variation of Verdoorn's model provides the basis for what are really two papers. One paper is an estimate of the potential increase in demand for manufactures from developing countries with a 50 percent preference. He should have assumed zero duty treatment, with exceptions or safeguards applied to the more sensitive products (i.e., those with high supply elasticities and high elasticities of substitution). Nonetheless, his figures fall in the same ball park as those of other economists as well as internal government estimates. The other paper within the paper is a clever manipulation of the Verdoorn model with the use of conventional, Slutsky-style utility theory. He makes assumptions rather freely, but they seem reasonable. What else can one do about export supply elasticities, or elasticities of substitution? What is of interest here is the theoretical generality of this model of the effects of discrimination. It is just a beginning, but it is worth further exploration. On the whole, however, I must con-

fess a strong desire to see these models constructed in terms of rates of growth of income and of prices, so that the dynamic of high import demand with inflation, and the dynamic of overall trade-payments adjustment, are built in. The earlier observations by Balassa and Truman, referred to by Clague, that the degree of trade diversion and trade creation depend on capacity utilization (full employment) seem right, and if they are, this requires analysis in the latter dynamic framework.

STEPHEN GUISENGER: In his paper Bela Balassa argues that the welfare maximizing strategy for a developing country is to set uniform effective rates of protection across all activities with exceptions being granted only for infant industries and optimal tariff situations. What Balassa documents in the case of the nine developing countries he has studied in depth over the last few years is that uniform effective rates of protection have not been observed in practice and that the costs of the trade policies pursued in these countries in terms of real income foregone have been considerable. Since I am in substantial agreement with Balassa on both his diagnosis of the effects of trade policies in developing countries and his recommended cure of more uniform tariff and exchange rate policies, this discussion will simply expand on Balassa's paper, first, by identifying several important assumptions which are implicit in Balassa's analysis and, second, by indicating where I believe Balassa's analysis leads us in terms of further empirical research on trade policies of developing countries.

One fundamental assumption extending throughout Balassa's paper is that trade policies of developing countries *do* make a difference in the allocation of resources within the economy. This may seem a patently obvious assumption to many, but much of the recent empirical research on developing countries, especially the cross-section studies of the patterns of industrialization, has emphasized factors other than trade policies in the determination of the levels of output in different economic sectors. These studies have stressed the importance of population size, per capita

income, technological efficiency and relative factor abundance in determining comparative advantage and industrial specialization. Balassa's analysis, based on detailed case studies of trade policies in nine developing countries, concludes that commercial policies in the form of tariffs, export taxes and subsidies, quotas, and multiple exchange rates can significantly distort the pattern of industry output and factor usage away from the free trade pattern of specialization. The assumption that trade policies are effective is, of course, necessary in order to maintain that differentiated rates of effective protection among industries have inflicted very considerable costs on developing countries.

I would like to mention briefly one other important assumption in Balassa's analysis which concerns the protection afforded export industries. Balassa assumes that exports deserve levels of protection equal to the levels afforded import substituting activities. I would like to add that there are situations where the preferential treatment of export activities is warranted. Staffan Burenstam Linder in his book *Trade and Trade Policy for Development* notes that exports can act as a "super engine" for growth, exerting a leverage effect on the economy many times greater than their direct contribution to GNP. If such a multiplier effect does in fact exist then preferential treatment would be called for, provided that only small increases in protection were required to expand the quantity of foreign exchange earned. The paper by Ingo Walter presented in this session suggests, however, that the price elasticity of demand in the developed countries for the exports from developing countries may in fact be relatively low, owing to the presence of nontariff barriers on products where developing countries have a strong competitive position.

Where does Balassa's analysis lead us then in terms of future empirical research on the trade problems of developing countries? Many recent empirical studies of systems of tariff protection have employed the concept of the effective rate of protection. Rates of effective protection can be quite different from the rates of nominal protection afforded the final products of industries. Balassa's diagnosis of

the inefficiencies of trade policies in developing countries has been based largely on his observation of wide disparities in effective rates among industries within these countries.

The actual measurement and interpretation of effective rates of protection remain very controversial matters, however, as was clearly demonstrated at the recent conference on effective protection organized by the GATT and the Graduate Institute of International Studies in Geneva. At the conference, questions were raised about the ability of a partial equilibrium formula for effective protection to accurately measure the resource allocative effects of a tariff system given that general equilibrium repercussions are ignored. Additional questions were raised concerning the implications of substitution among inputs in the production process and of less than perfect elasticity in the foreign demand and supply of commodities for the measurement of levels of effective protection. Because of these and similar questions more than just a few of the participants at the conference held the view that the rates of effective protection calculated by the conventional formula are totally meaningless, if not even slightly misleading. Thus, to some economists, Balassa's conclusions regarding trade policies in developing countries, which are reached on the basis of the observed levels of effective protection, remain to be proven for lack of an acceptable indicator of tariff protection.

In my view the next step in the empirical analysis of trade policies for developing countries must be to validate the conventional measure of the effective rate as the appropriate indicator of the resource allocative effects of a tariff system. One aspect of this validation process would focus on the sensitivity of both absolute and relative levels of effective protection to changes in the assumptions underlying the conventional formula for effective protection while the second aspect would demonstrate that changes in the allocation of resources in developing countries can indeed be traced to changes in levels of effective protection. This last aspect is particularly difficult because movements of factors among industries are rarely due to only one cause and the statistical tests for this relationship will have

to be carefully worked out in order to isolate the independent effect of protection on factor movements. Nevertheless, it would seem to me that the statistical validation of the causal link between protection and resource movements is imperative if analyses of trade policies using measures of effective protection are to assure their credibility.

One final area in which further empirical research is necessary is in the characteristics of protected industries in developing countries. We still do not know, for example, whether tariff structures in these countries tend to protect capital, labor or some other input. We do not know whether protection has seriously lessened competition within the economy and led to significant losses in production efficiency. We know very little about the income distribution effects of tariff systems. The several studies of trade policies conducted by Balassa provide considerable insight into these questions, but we need much more information on all of the above questions before we can begin to implement the kind of policy guidelines sketched out by Balassa in the conclusion to his paper. Tariff setting in any country is a political, not a market, process and we can gain an understanding of how this political process has operated in the past by studying carefully the types of industries which have received tariff protection. Armed with an understanding of the political process at work in each country, we may be able to offer a developing country more than just the recommendation that it move towards uniform effective rates of protection; we may in fact be able to offer it a politically feasible plan for tariff reform that will achieve that end.

ROBERT E. BALDWIN: As more and more quantitative knowledge has been accumulated concerning the trade policies of both the developed and developing nations towards each other, we have increasingly come to appreciate the great difficulties that the developing countries must overcome if they are to increase substantially their exports to the developed nations. A few years ago Bela Balassa brought out the fact that the levels of nominal and, especially, effective tariff rates imposed by the industrial nations against manufactures im-

ported from developing nations were significantly higher than those applying to all imports of manufactures. Now Ingo Walter in his interesting paper has presented data indicating that this discrimination against the developing nations is further increased through the use of various nontariff policies by the developed countries. Thus far, the data assembled relate to the extent of trade subject to nontariff barriers without reference to the degree of restrictiveness of the various barriers, but there seems little reason to believe that the conclusions will change when more extensive quantitative estimates of the impact of NTB's are available.

My own analysis of the nontariff barriers imposed by the United States tends to confirm Walter's results. Taking into account the main quotas, export subsidies, and federal, state and local indirect taxes, my estimate of the average rate of effective protection in the U.S. due to nontariff barriers alone is about 7 percent. However, when the effective levels of protection by industry are weighted by imports from the LDCs rather than by imports from all sources the average effective rate increases to 19 percent. Clearly, such measures as quotas on sugar, oil, and cotton textiles fall with particular force on the exports of the LDCs. The threatened quantitative restrictions on shoes and wool and man-made fibers will further increase the discrimination against the exports of these countries.

In the face of the strong movement over the last twenty-five years toward trade liberalization, generally only relatively depressed industries as well as sectors supposedly particularly important for national defense purposes have been able to increase their protective levels or at least resist any significant cuts in these levels. Prominent within this group of industries have been those in which the LDCs already have or are developing a comparative cost advantage, namely certain natural-resource industries and industries employing relatively large numbers of unskilled labor. Moreover, the protection given to these industries has increasingly taken the form of nontariff measures. In this regard Professor Walter's point about the greater restrictive effect of any given nontariff trade-distorting measure on a

developing country in contrast to a developed economy deserves special emphasis. Exporters who wish to sell their products in markets protected by nontariff trade barriers generally are forced to incur considerably greater cost outlays in finding out what the exact nature of the barriers are than if the protection takes the form of tariffs. Because of the rather significant scale economies associated with obtaining information, this means that new and small exporters of manufactures—like most developing countries—are placed at a substantial disadvantage relative to large established producers.

The task of increasing exports by the developing countries is made especially difficult not only because the products of export interest to them are subject to disproportionately high tariff and nontariff barriers, but also because the world markets for these products tend to grow less rapidly than for world exports as a whole. Indeed it is in large part because these items are in relatively declining commodity groups that the incidence of tariff and nontariff barriers levied against them are relatively high. Imports of the countries to which developing nations send most of their exports are also not growing as rapidly as world imports. If the constant market-share technique is used to estimate the effect of changes in the commodity and country composition of world trade on the export share of the LDCs in world trade, the results indicate that between both 1961–65 and 1965–69 these two effects more than accounted for the decline in the LDC's export share of all world trade as well as their share in world trade of manufactures. This means that a positive "competitive" effect was more than offset by unfavorable structural shifts in world trade.

A fourth set of factors helping to explain

the difficulties faced by the developing countries in increasing their exports are those discussed by Bela Balassa in his very rich paper, namely, the trade policies of the developing countries themselves. One point he raises that I would like to comment upon concerns the issue of whether, aside from policies designed to correct cost disadvantages or take into account infant-industry grounds for protection, manufacturing should receive preferential treatment. He concludes that there is some presumption in favor of promoting the manufacturing sector because of labor training and interindustry linkage effects. I would stress the case for manufacturing even more strongly. In evaluating alternative growth strategies in the developing nations we must always take account of the cost of implementing social and economic objectives relating to such matters as income distribution, employment, land reform, the spreading of technological and market knowledge, and political modernization. Economists too often tend to regard these costs as independent of the development policies followed. Thus the strategy that maximizes the short-run increase in real income seems best. In fact, however, the economic and political difficulties involved in directly promoting these objectives in developing countries are often so high that the spillover effects of particular development strategies in furthering them become a major factor influencing the decisions of policy makers. While lack of time prevents me from elaborating upon the theme, I think that the case for providing the type of preferential treatment to manufacturing outlined by Balassa can be further strengthened on the grounds that it also best promotes some key broad economic and social objectives of the developing nations.

EVALUATION OF ECONOMIC REGULATION OF INDUSTRY

Methodology of Evaluating Economic Regulation*

By FRED M. WESTFIELD
Vanderbilt University

One of the principal quantitative techniques used in the study of costs and benefits of regulation involves the evaluation of consumer's and producer's surplus. Examples are [2], [3], [4], [6]. The method can be caricatured as follows: Obtain data on quantity and price for the output produced by the regulated industry. Obtain estimates of the slopes of the demand and marginal cost from cross-section or time-series estimates of elasticities. Calculate where demand and marginal cost would intersect. Complete a triangle whose vertices are (1) the predicted intersection of marginal cost and demand, (2) the current demand price, and (3) the current estimated marginal cost. Add to the area of this dollar triangle the direct total cost of the entire regulatory machinery—the budget of the regulatory agency and the budget for lawyers, accountants, engineers, public relations men, and for all the other costs incurred by those being regulated—and you have an estimate of the social cost of regulation. Waiving their validity and accuracy for the moment, the calculations assume that a feasible social alternative to the regulated status quo is, in fact, described by the intersection of the measured marginal cost and demand curves, and that this alternative situation can be

reached without new direct regulation costs offsetting the savings.

If the social optimum were to require a price-output configuration for the regulated industry described by the price-marginal cost equality, and this equality could be brought into being by a costless restructuring of the regulated industry into one behaving like a competitive industry, then the social choice is trivial. The complex statistical calculations are unnecessary. However, if pure competition or its simulation are not viable alternatives to the regulated status quo because of decreasing unit costs often considered characteristic of public utilities, then the calculations are beside the point. Dismantling the regulatory machinery will save the resources used by the regulatory process, but the *laissez-faire* outcome will probably be oligopoly with its absence of price competition, excessive product differentiation, wasteful sales promotion and advertising, excess capacity, and expensive legal talent to forestall and defend antitrust prosecutions. The measured demand and marginal cost curves and the triangle provide no information about a new deregulated equilibrium. And it is not a valid proposition that entry of firms, threats of entry, and oligopolistic rivalry will be an improvement over regulated monopoly.

So far I have not challenged the proposition that the intersection of demand and

* Research support of the National Science Foundation is gratefully acknowledged.

marginal cost curves estimated from current and historical data reveal the price and output required for the social optimum. Actually the intersections of such observed demand and marginal cost curves are of no normative social significance at all. Other prices in the economy that determine the positions and shapes of both marginal cost curve and demand curve of the regulated industry are going to change because of income and substitution effects when regulated price and output are altered. Furthermore, other prices are generally not determined by their marginal cost-demand curve intersections. Even casual observation of the structure and behavior of supplier markets for regulated industries reveals that input prices and the prices of inputs to inputs have substantial wedges between them and their marginal costs because of noncompetitive market structures, if not because of tariffs, quotas, taxes, and the like. Choices of efficient input combinations are distorted and pyramided into the cost function of the regulated firm. Indeed, as will be emphasized below, regulatory processes themselves cause distortions of costs. The demand data observed for a regulated industry are also invalid for normative analysis. For example, the output often is an intermediate good embodied in final goods sold at prices higher than their marginal cost. Derived demand curves faced by the regulated industry and estimated for it thus are marginal revenue products less than the value-of-marginal-products relevant for efficiency.

Even if we pretend that the above considerations do not invalidate the measurement of the triangles, there are other difficulties with the mode of analysis. Simultaneous social marginal cost-demand intersections everywhere actually do not characterize a social optimum unless, as a further condition, income is also distributed optimally. Demand curves (and marginal

cost curves also) obviously will shift with alterations in the distribution of income among people with differing income elasticities of demand. Which of the marginal cost and demand curves would be appropriate for the calculation of surplus triangles? Each set of marginal cost-demand intersections, efficient in the Pareto sense, generally gives a different solution for prices and outputs. One might be justified in supposing that the income distribution problem is of no concern for the evaluation of regulatory processes if there were practical methods by which incomes could be redistributed using lump-sum taxes and transfers, methods which would not themselves create price-marginal cost wedges and inefficiencies of substitution effects. The optimal demand and marginal cost intersection points are those generated by an income distribution which those prices and lump-sum transfers, made by policy makers charged with optimizing distribution of income, bring about. The truth is that lump-sum transfers are not practicable policy instruments. Income taxes, sales taxes, antitrust enforcement, public utility regulations, public goods, and all the other instruments of government influencing resource allocation tend to leave wedges with substitution effects causing inefficiency, at the same time that they create the income effects redistributing welfare. The social optimum has to be the result of tradeoff between efficiency and fairness, and it is generally not given by supply-demand intersections. As Samuelson [5] has summarized the proposition "... it is a theorem in *optimal-feasible* welfare theory . . . that *some* departure from Pareto-optimality is optimal in the interest of interpersonal optimisation."

I concluded earlier that calculation of the areas of surplus triangles was not really useful. The additional discussion of the properties of the social optimum strengthens the conclusion. Calculations of the

dollar magnitudes of areas of triangles are probably harmful. They make it appear as if questions about costs and benefits of regulation are or could be resolved scientifically, independently of one's views about the deservingness of claims to national product by the contesting groups in the society.

* * *

Some of the issues that I have raised quite generally are more fully illuminated by closer analysis of a generalized Averch-Johnson [1] model of the regulated monopoly firm.

The model is that of a profit-maximizing monopolist given an exclusive franchise, whose rate of return on capital is limited to a value, say r^* , by a regulatory commission. The value r^* is assumed to lie between the level r^0 , associated with profit maximization, and \bar{r} , the prevailing cost of capital. These restrictions insure that regulation is effective in truly constraining the firm to earning a reduced rate of return without forcing it to incur unsustainable losses. The system of regulation causes output and substitution effects that are exactly the same as those of a subsidy or tax reduction based on value of capital services purchased by an unconstrained monopolist, and the direct opposite of a tax levied on the firm's earnings on capital. Of course, the income distribution consequences are quite different.

Equilibrium conditions for the constrained monopoly are quickly summarized when $\bar{r} < r^* < r^0$. Suppose there are v non-capital inputs described by the vector V and k capital inputs by the vector K . There is a production function for the regulated output $q \leq F(V, K)$ whose marginal physical products for non-capital input are F_i and for capital inputs F_{v+j} . Generalized diminishing returns are assumed, with the exception that there may be increasing returns to scale. Marginal

input costs of noncapital inputs are n_i and their prices are w_i . Marginal input costs of the capital goods are m_j and their prices u_j , thus the marginal costs of the services from the non-depreciating capital goods are $\bar{r}m_j$ and their average costs are $\bar{r}u_j$, where \bar{r} is the rate of interest. If the firm is a price-taker with respect to inputs, $n_i = w_i$ and $m_j = u_j$. The downward-sloping demand curve for the output sold by the regulated firm is summarized by $p = p(q)$, where p is the average price of the regulated output.

From profit maximization one obtains the first-order conditions:

$$(1) \quad F_i(V, K)/n_i = 1/MC(q | uK) = 1/MR, \\ i = 1, 2, \dots, v,$$

$$(2) \quad F_{v+j}(V, K)/(\bar{r}m_j) = 1/MC(q | wV) \\ < 1/MR, \\ j = 1, 2, \dots, k,$$

$$(3) \quad \left(p(q)q - \sum w_i V_i \right) / \sum u_j K_j = r^*.$$

Equation set (1) says that noncapital goods are purchased so that their marginal revenue products are equal to their marginal input costs. Output q must be selected so that the "short-run" marginal cost obtained by varying the noncapital inputs only (holding capital inputs constant), and here designated by the function $MC(q | uK)$, is just equal to the marginal revenue MR from output sold. Equation set (2) shows that the capital inputs, constant in set (1) must be allocated so that, for a given capital outlay $uK = \sum u_j K_j$ and for given quantities of non-capital inputs V , output is maximized. The set of equations (2) determines the optimal combinations of the capital goods to be held constant in the set (1), just as the set (1) determines the optimal combinations of the noncapital goods held constant in the set (2). The set (2) also

shows that total capital expenditures are greater than those which would be required in the absence of an effective rate-of-return restriction on the monopolist. The equal marginal productivities per incremental dollar of capital services in (2) are lower than the equal marginal productivities per incremental dollar of the noncapital goods expenditures in (1). Equation (3) shows by how much capital expenditures have to exceed the amount required for unconstrained profit-maximum output. The least-cost mix of capital expenditures must be increased so as to just satisfy together with the reoptimized mix of noncapital inputs the rate of return restriction (3).

By formulating the problem with input prices differing from marginal input cost and allowing for monopsony, I emphasize that lack of proportionality between marginal productivities and input prices as a result of a rate-of-return constraint and the resulting cost curves observed by economists are of no significance for welfare analysis. I will now continue the analysis by assuming for simplicity that the firm is a price-taker. There is still, of course, no presumption that prices of inputs correspond to social marginal costs of producing them.

Important aspects of the model can be illustrated on a conventional monopoly price-output diagram. In Figure 1, the unconstrained long-run monopoly equilibrium is given by the output q^0 and price p^0 determined by the intersection of the conventional long-run marginal cost $MC(q)$ and marginal revenue curves $MR(q)$. Also passing through the intersection point of the MC and MR curves is a constrained marginal cost curve $MC(q|uK^0)$ which shows how marginal cost would vary if output varies as noncapital inputs are adjusted so as to minimize their costs, while the dollar value of the capital stock, but not its composition, is held con-

stant at its unregulated long-run equilibrium value, say $uK^0 = \sum u_j K_j^0$. This marginal cost curve is like a short-run marginal cost curve. However, instead of holding physical capital inputs fixed, as a short-run curve usually does, here it is assumed that there is a fixed *capital outlay* and that the various physical capitals, whose services are purchased at prices exogenously given to the firm, are efficiently reallocated so as to maximize output as output changes. In its unconstrained long-run equilibrium position, the firm is hypothesized to earn a rate of return r^0 on capital stock uK^0 which is in excess of the rate of return r^* allowed by the regulatory commission. Profits are too high relative to the value of the capital stock uK^0 .

Of all the methods open to the firm for bringing profit into line with the value of the capital stock, the generalized Averch-Johnson result requires (a) increasing the dollar value uK^0 of the capital stock to a larger value uK^* and (b) adjusting inputs so that the marginal cost of output resulting from the variation of the noncapital inputs alone, with capital expenditures held constant at the higher level but allo-

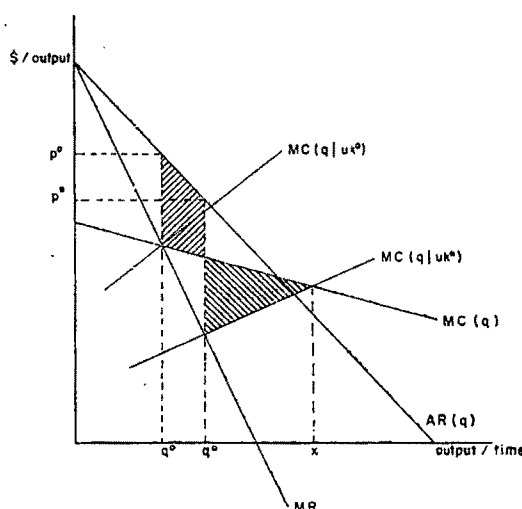


FIGURE 1

cated among various capital goods efficiently so as to maximize output, is just equal to the marginal revenue from the output. Increases in the value of the capital stock and the reallocations of the capital inputs from the set K^0 to K^* and of the noncapital inputs from the set V^0 to V^* , as output is adjusted from q^0 to q^* , must continue until the permitted rate of return r^* is just equaled by the ratio of revenue from sales (p^*q^*) net of total noncapital cost (wV^*), to the value of the capital stock (uK^*), i.e., Equation (3). The increases in capital expenditures required to meet the rate-of-return ceiling will lead to a shift in the "short-run" marginal cost curve that shows how marginal cost of output varies with a fixed total of optimally allocated capital expenditures. Thus, in Figure 1 the intersection point of the "short-run" marginal cost curve $MC(q|uK^*)$ with MR gives the equilibrium output q^* , and the corresponding price p^* on the demand curve gives the price for the regulated profit maximizing monopolist permitted to earn r^* . Outlays among the various noncapital inputs and among the various capital inputs are least cost from the point of view of the firm, at the same time that the equal marginal products per dollar for all of the noncapital inputs exceed the equal marginal products per dollar for all of the capital goods' services, so that output is not produced at least cost because of overspending on the total of capital goods and underspending on the total of noncapital goods, i.e., Equations (1) and (2). This subtle inflation of cost would probably escape econometricians measuring waste as an area between demand and cost curves.

There has been some discussion by theoretical economists on whether imposition of a rate-of-return ceiling necessarily increases equilibrium output and lowers price of the profit maximizing monopolist. That imposition of the rate-

of-return ceiling leads to additional capital outlays is not controversial: $uK^* > uK^0$ because of the second-order conditions for the maximum. However, might not the increases in capital outlays lead to decreases in the noncapital outlays to an extent that profit maximizing equilibrium output of the regulated monopolist decreases, so that price actually increases with imposition of (or lowering of) the allowed rate of return? Figure 1 indicates the answer to this puzzle. If the "short-run" marginal cost curve shifts down and to the right as capital is increased from uK^0 to uK^* to satisfy the regulatory requirement, then regulated equilibrium output q^* must obviously be greater than unregulated equilibrium output q^0 , and regulated equilibrium price p^* must be less than unregulated price p^0 . Under what conditions will the increase in capital now lower the "short-run" marginal cost curve? Since relative prices among capital inputs and among noncapital inputs are unchanging, we can think of this as a two-input problem—two Hicksian composite inputs. And it has been proven rigorously [7, p. 430] that the short-run marginal cost curve will shift down and to the right provided the composite capital input (uK^*) is not an "inferior" input. (An inferior input is one whose quantity demanded decreases with increases in the output of a firm while marginal technical rates of substitution are constant—a bending back of the expansion path.) Equivalently, because of duality, if a proportional fall in price of all capital inputs of the *unconstrained* profit-maximizing monopolist would always raise equilibrium output and lower its price, then imposition of (or lowering) an existing rate-of-return ceiling on such a monopolist will also surely raise output and lower its price. For public utilities it seems empirically quite unlikely that the composite capital input would be inferior. Therefore, the imposition of the restriction on rate of

return (or its tightening) generally will increase equilibrium output and lower the price.

By how much are private costs increased as a result of the distortion from rate-of-return regulation? Figure 1 allows one to read off this amount. Consider the intersection point of the long-run marginal cost curve, $MC(q)$, with the "short-run" marginal cost curve, $MC(q|uK^*)$, at an output, say x . This indicates that the capital outlay $uK^* = \sum u_j K_j^*$ used in excess because of rate of return regulation in the production of the output q^* , is exactly the capital that would be utilized by an unconstrained firm in producing an output x efficiently in long-run equilibrium. In other words, uK^* is the capital outlay, allocated among the various capital goods so as to maximize output, which together with appropriate amounts of noncapital inputs, would be the ideal amount of capital for the production of the larger output x . The costs of additional (noncapital) inputs that would be required to produce the output x , for which the capital uK^* is ideal, instead of the regulated output q^* , for which this capital is excessive, is given by the trapezoidal area under the "short-run" marginal cost curve for the fixed capital outlay uK^* , i.e., $MC(q|uK^*)$, between the outputs q^* and x . On the other hand, the costs of additional inputs that would be required to produce the output x instead of the regulated output q^* , if the output q^* were *not* produced inefficiently with the excessive capital, but were produced instead efficiently with the least-cost combinations of capital and noncapital inputs—this cost is given by the trapezoidal area under the long-run marginal cost curve $MC(q)$, between the outputs q^* and x . Thus the difference between these two areas—the shaded triangular area between the long-run and "short-run" marginal cost curves that lies between output q^* and x —is an exact

measure of the excess private costs incurred as a result of producing regulated output q^* with the excessive capital induced by rate-of-return regulation rather than with a least-cost input mix.

But I must insist that the triangle measures only private cost. These private costs may even be public benefits. Imagine that a corporation income tax had raised the effective cost of capital to the firm and reduced its capital intensity. Might not the triangle as plausibly be a measure of social benefit from rate-of-return regulation that restores tax-distorted capital/noncapital ratios to more appropriate levels?

Having discovered such a cost triangle, a true believer in the consumers' surplus measurement-game would not be able to resist the temptation to relate it to the shaded "benefit" trapezoid on Figure 1 lying between the long-run marginal cost curve and the demand curve as output is increased from q^0 to q^* . Indeed, since both the benefit trapezoid and the cost triangle increase as the allowed rate of return r^* is lowered and the "short-run" marginal cost curve, $MC(q|uK^*)$, shifts down, he would find that regulated rate of return, r^* , which maximizes the difference between benefit and cost. I have said enough in this paper already on why I disapprove. Observe also that there are feasible regulatory techniques which could induce the firm to produce the output q^* with a smaller or indeed zero cost triangle. [8]

The model of the monopoly firm analyzed here reveals one further problem for regulation and also for the measurement of its effects that requires more attention of economists. This arises with inelastic average revenue curves. Suppose that a rate-of-return regulated, profit-maximizing monopolist has adapted his capital stock so that the intersection of the "short-run" marginal cost $MC(q|uK)$ with MR occurs in Figure 1 on the hori-

zontal axis, and the firm still realizes receipts net of noncapital costs that are greater in relation to the capital base than the regulatory authority has permitted. Lowering price will increase output, and because demand is incipiently inelastic, the declining total receipts, possibly accompanied by further growth in the capital base, will insure that at some price the permitted rate of return will no longer be exceeded. However, if the firm does not lower price but instead increases the capital base without increasing quantity sold, it will also be able to meet the rate-of-return ceiling and do so with larger total profits. With demand inelastic, cost-based regulatory techniques such as rate-of-return or operating-margin restrictions fail to bring price and output responses. Costs are inflated, firms are off their production functions, and the resources brought into the industry in response to the regulatory restrictions have zero marginal products or worse. Ironically, attempts at evaluating areas of surplus triangles will probably fail to detect this nearest approximation to pure waste because statistical cost-output data will not reveal the excess costs of inputs.

In this paper I have emphasized difficulties of measuring costs and benefits of regulation. Furthermore, I have tried to argue that total dollar figures are unnecessary, beside the point, or worse. The real-world choices are generally not between regulation and no regulation, but between various methods of regulation and degrees

of regulation. They are between alternative means of achieving objectives or marginal changes in objectives. Research on how alternative methods and degrees of regulation influence behavior of regulated industries and the economy, and on how they affect income distribution, can provide critical information for the public policy choices to be made.

REFERENCES

1. H. Averch and L. L. Johnson, "Behavior of the Firm Under Regulatory Constraint," *A.E.R.*, December 1962.
2. F. Darby, *An Evaluation of Federal Regulation of Common Motor Carriage* (Unpublished Indiana University Ph.D. dissertation), 1969.
3. A. Friedlaender, *The Dilemma of Freight Transport Regulation*, (Brookings, 1969).
4. P. W. MacAvoy, "The Effectiveness of the Federal Power Commission," *Bell Journal of Economics and Management Science*, Autumn 1970.
5. P. A. Samuelson, Review of J. E. Meade, *Efficiency, Equity and the Ownership of Property* in *E.J.*, December 1965.
6. R. M. Spann and E. W. Erickson, "The Economics of Railroading: The Beginning of Cartelization and Regulation," *Bell Journal of Economics and Management Science*, Autumn 1970.
7. F. M. Westfield, "Regulation and Conspiracy," *A.E.R.*, June 1965.
8. F. M. Westfield, "Innovation and Monopoly Regulation," Chapter 2 of W. Capron, ed., *Technological Change in Regulated Industries*, (Brookings, 1970).

Costs and Benefits of Regulating Communications

By BOYD L. NELSON*

Federal Communications Commission

It has been said that "there is no general theory of public utility regulation."¹ This statement may be a subtle way of implying that the regulators of public utilities don't know what they are doing, or at least do not have a general plan or set of guidelines in the form of a general theory. However, I think it is the case that most regulators, although there are exceptions, do know what they are doing and know even better what they would like to do if the resources were forthcoming to allow them to do it. Furthermore, there is no doubt about the fact that there exists a generally understood rationale of public utility regulation which I think can be summed up rather succinctly.

Suppose a desirable level of demand for a service basic to the efficient functioning of the entire economy is given. If the structure of the producing industry is such that the results in terms of service quality and price are highly indeterminate, or if the structure is such that the results are determinate but considered bad in the light of public policy and social welfare goals, then we may decide to regulate by way of commission authority. Under commission regulation, we essentially seek to require the utility companies to operate at cost, where the word "cost" includes the cost of capital. The desired level of cost is given

by the following algebraic expression: $C = E + r(P - D)$. In this relation, C stands for total annual cost at some given point in time, E stands for operating expenditures including taxes but excluding interest, and r denotes the cost of capital, this being an average figure covering both debt and equity; P is the book value of the gross plant devoted to the provision of utility services, and D is the value of the depreciation reserve. C represents an optimal condition. That is, given the level of demand and the quality constraints, C has a minimum value.²

A statement of the corresponding revenue relation is as follows:

$$R = \sum_{i=1}^k q_i p_i,$$

wherein it is assumed that k different services are being offered, with q_i being the quantity of the i^{th} service and p_i its price. The expression is, of course, an oversimplification, for the rate structure of each service can be rather complex, but by subdividing each of the service structures and using several summations, we could devise an expression which closely represented the actual situation. The prices are to be set so that $R=C$ and acceptable anti-discriminatory constraints are met. We shall call the value of R the revenue requirement and that of C the cost requirement. We shall use the symbol R' to denote the actual revenue inflow. The public in-

*The author is Chief of the Economic Studies Division, Common Carrier Bureau, FCC. The views and opinions expressed herein are those of the author and do not necessarily coincide with those of the Federal Communications Commission or other members of its staff.

¹ Richard Gabel, p. 340.

² Regardless of pricing methods or the use of subsidies, this cost must be met.

terest is best served if the revenue inflow equals the cost requirement ($R' = C$), and it is the business of the regulatory authorities to bring the former down to the latter in the face of ongoing tendencies that give rise to differences between the two values. One such tendency is the improvement in efficiency of operations arising from technological progress over time coupled with the inevitable regulatory lag. Another one is the continual exertion of monopoly power by the utility and the virtual impossibility of complete regulatory control of such power.

The relations set forth for C and R point to four particularly important problems to be attacked in public utility regulation, namely (1) the estimation of the probable level and pattern of demand and the values of the relevant demand elasticities, (2) the estimation of the cost of capital, (3) the determination of proper constraints on price discrimination among services and among users, and (4) the estimation of appropriate depreciation rates and the formulation of appropriate depreciation methods. If these four problems could be solved satisfactorily, the remaining regulatory problems would be relatively minor and have to do largely with expense audits and other accounting functions of a surveillance character. In an age of electronic computers and sophisticated statistical sampling techniques, it would seem that these problems could be handled with facility. Let us, then, indicate the scope and limits of our investigation.

We shall be concerned here only with *ex post* benefits and costs relating to common carrier communications as they are regulated at the federal level. This limitation will confine us to interstate toll and overseas telephone calls and record messages, the latter including both data and other forms of record communications. The period of the investigation shall be from fiscal 1951 through fiscal 1970. The Com-

mon Carrier Bureau of the Commission did not exist prior to April 1950, and only after that date can a satisfactory separation be made of money spent exclusively on common carrier regulation.

Within the confines we have just set forth, we must next define what we mean by costs and benefits. We shall define costs initially as those portions of the FCC appropriations which are obligated for common carrier regulation. We shall define benefits as the effects on the aggregate revenue inflow arising from rate changes in each fiscal year. In general, the benefit is considered positive if the revenue inflow decreases; the benefit is considered negative if revenue inflow increases. The built-in regulatory lag, which has been mentioned previously, leaves the effect on revenue inflow behind the times, so to speak, so that as a practical matter the revenue inflow will almost never be equal to or less than the cost requirement. Legal procedures take a certain necessary amount of time, and apparently this is something that cannot effectively be reduced by management science or any other technique known to man.

We emphasize that we are confined by jurisdictional considerations to the interstate toll and United States-overseas types of communications. We shall have more to say about the effects of this limitation later. So far, then, we could take as costs the flow of tax money which is channeled into the obligations for common carrier communications regulation. And we could take as benefits the cumulative amount of reductions in revenue inflow that resulted over the years from various rate changes in the communications field. Utilizing these two unadjusted amounts we could construct a crude benefit-cost ratio. Benefit and cost flows of these types have been tabulated in Table 1 and the crude benefit-cost ratio computed. We note that most of the rate reductions have been in-

TABLE 1—CRUDE COSTS AND BENEFITS
(Thousands of Dollars)

Fiscal Year	Costs	Benefits		
	Obligations for Common Carrier Regulation	A.T.& T.	W.U.*	International Record Carriers
1951	866	92	(1,361)	
52	823	(14,435)	(10,739)	
53	761		(12,217)	(110)
54	773	(73,350)	(357)	
55	800	262	(10,302)	
56	818			
57	886	477	(9,950)	
58	1,009		(137)	
59	1,089	(2,427)	(13,223)	(3,094)
60	1,212	50,500	(65)	
61	1,347	322	(10,069)	
62	1,352		191	
63	1,646	30,000	(5,255)	
64	1,901		(8,322)	
65	2,109	102,832	(4,490)	
66	2,156		(473)	(1,779)
67	2,223	23,749		
68	2,218	115,226	(13,005)	10,500
69	2,436	(44,863)	(9,483)	
70	2,687	160,150	(19,404)	5,486
	29,112	348,535	(128,661)	11,003

* Includes some intrastate effects.

Grand Total of Revenue Effects: 230,877

Crude Benefit-Cost Ratio: $230,877/29,112 = 7.931$

curred in the telephone area and most of the increases in the telegraph area. The grand total of revenue effects (or benefits) is about \$231 million, and the grand total of costs over the same period is about \$29 million, making a crude benefit-cost ratio of about 7.9. The cost of common carrier regulation in 1951 incidentally was \$866,000, and in 1970 \$2,687,000. There are at present about 150 people concerned with common carrier regulation at the federal level in the communications field, and this number includes all kinds of job classifications from clerk-typist trainee on up to the highest paid lawyer. Most revenue inflow reductions deriving from rate cuts in the telephone field have occurred since 1959. (We are probably fortunate that the study period ends in fiscal 1970.)

There are some fairly obvious adjustments that must be made to the cost figures of Table 1 so that they will relate more closely to the benefits we are measuring. On the one hand, the cost of legal proceedings incurred by the carriers should be included, and on the other hand, costs which do not relate to benefits as we have defined them should be deducted. The former amounts have been compiled from reports of the carriers and the latter have been estimated by administrative agency personnel, such estimates covering the costs of licensing point-to-point microwave stations and land-mobile radio stations. The adjusted costs appear in Table 1A and the revised benefit-cost ratio has been computed thereon at 7.6.

However, there are still several difficulties connected with the figures entering

TABLE 1A—COSTS AND BENEFITS
(Thousands of Dollars)

Fiscal Year	Costs	Benefits		
	Adjusted Obligations for Common Carrier Regulation	A.T.& T.	W.U.*	International Record Carriers
1951	712	92	(1,361)	
52	622	(14,435)	(10,739)	
53	596		(12,217)	(110)
54	598	(73,350)	(357)	
55	605	262	(10,302)	
56	640			
57	754	477	(9,950)	
58	825		(137)	
59	870	(2,427)	(13,223)	(3,094)
60	962	50,500	(65)	
61	1,073	322	(10,069)	
62	1,225		191	
63	1,340	30,000	(5,255)	
64	1,536		(8,322)	
65	3,013	102,382	(4,490)	
66	3,088		(473)	(1,779)
67	2,986	23,749		
68	2,630	115,226	(13,005)	10,500
69	3,039	(44,863)	(9,483)	
70	3,215†	160,150	(19,404)	5,486
	30,329	348,535	(128,661)	11,003

*Includes some intrastate effects.

† Estimated in part.

Grand Total of Revenue Effects: 230,877

Crude Benefit-Cost Ratio: $230,877/30,329 = 7.612$

this last computation. We shall set out to examine some of them and after so doing, attempt to incorporate appropriate adjustments into the analysis and note the effect on the value of the ratio.

The first difficulty involves the time value of money. In the crude *ex post* streams of costs and benefits of Table 1A, there has been no adjustment for this factor. We might say that if tax monies had been saved in the event of no regulation, then such amounts could have been accumulated at interest over time. Similarly, in the case of benefits, the amount of benefit, or expenditure reduction for users of the service, could have been invested and accumulated at interest. Utilizing these concepts, we find that the costs are somewhat greater and the benefits somewhat less

than in the case of the crude benefit-cost ratio. Difficulties arise concerning the appropriate rate of interest to use, but we have chosen 4% as a rate which is a reasonable average, we think, for low risk investments over the period under consideration. Table 2 reflects the adjustment for the time value of money, which has the effect of reducing the benefit-cost ratio to a value of 4.1.

Secondly, we might consider adjustments for changes in the price level over time. In terms of goods and services foregone in tax payments and in terms of goods and services that are equivalent to the benefits forthcoming, the benefit-cost ratio might be considerably different from what we have computed in Table 2. Such adjustments can be made if we are willing

TABLE 2—TABLE 1A ADJUSTED FOR TIME VALUE OF MONEY AT 4%
(Thousands of Dollars)

Fiscal Year	Number of Years	Costs	Benefits		
		Obligations for Common Carrier Regulation	A.T.&T.	W.U.	International Record Carriers
1951	20	1,560.2	201.6	(2,982.1)	—
52	19	1,311.4	(30,412.4)	(22,625.5)	—
53	18	1,207.0	—	(24,749.4)	(222.8)
54	17	1,164.6	(142,878.5)	(695.4)	—
55	16	1,132.3	490.7	(19,295.5)	—
56	15	1,154.2	—	—	—
57	14	1,304.7	826.0	(17,230.0)	—
58	13	1,372.5	—	(228.1)	—
59	12	1,393.2	(3,885.7)	(21,170.4)	(4,953.6)
60	11	1,481.0	77,742.4	(100.1)	—
61	10	1,588.1	476.6	(14,904.6)	—
62	9	1,743.2	—	271.9	—
63	8	1,834.9	41,057.1	(7,191.8)	—
64	7	2,021.2	—	(10,951.2)	—
65	6	3,812.4	130,115.3	(5,681.3)	—
66	5	3,757.2	—	(575.5)	(2,164.4)
67	4	3,494.0	27,783.0	—	—
68	3	2,959.0	129,613.6	(14,628.9)	11,811.1
69	2	3,287.4	(48,523.8)	(10,256.8)	—
70	1	3,343.9	166,556.0	(20,180.2)	5,705.4
		40,922.4	349,161.9	(193,174.9)	10,175.7

Total Benefit = 166,162.7

Benefit-Cost Ratio = 166,162.7/40,922.4 = 4.060

to make assumptions about the kinds of goods and services involved. Assuming that wage goods or consumer goods are appropriate, we can use the Consumer Price Index in making adjustments. Table 3 incorporates such adjustments, with a resulting benefit-cost ratio of 2.2.

A question might arise as to whether the effects of rate increases should be counted as negative benefits. We feel that they should be so considered just as long as the estimated value of r in the cost formula is equal to or greater than the actual cost of capital. If however a situation should arise where the estimated value of r were less than the actual cost of capital, then rate increases could be a benefit so long as other expenditures (E), the rate of investment, depreciation rates, and service quality were optimal.

We are also faced with the question of whether benefits are reduced when a given rate reduction fails to bring the actual rate of return down to the cost of capital.³ It would seem that the benefit is not as large as it should have been, but perhaps the solution is to put more resources into the regulatory effort. Furthermore, in this particular day and age applied microeconomics is just now arriving at a threshold which indicates the feasibility of accurate estimates of the cost of capital.

Next, we have to face the question of the duration of the time period over which a given benefit lasts. In other words, we must consider the question, "when does the benefit from a given rate reduction

³ It should be noted that no unjustified diversions of benefits from customers to stockholders take place as long as the cost of capital is covered.

TABLE 3—TABLE 2 ADJUSTED FOR CHANGES IN THE PRICE LEVEL
(Thousands of Dollars at 1957-59 Prices)

Fiscal Year	Costs	Benefits		
	Obligations for Common Carrier Regulation	A.T.&T.	W.U.	International Record Carriers
1951	1,723.9	222.8	(3,295.1)	—
52	1,405.9	(32,878.3)	(24,460.0)	—
53	1,295.1	—	(26,555.2)	(239.1)
54	1,244.2	(152,648.0)	(742.9)	—
55	1,213.6	525.9	(20,681.1)	—
56	1,218.8	—	—	—
57	1,331.4	842.9	(17,581.6)	—
58	1,363.0	—	(226.5)	—
59	1,372.7	(3,828.3)	(20,857.5)	(4,880.4)
60	1,436.5	75,404.9	(97.1)	—
61	1,524.1	457.4	(14,303.8)	—
62	1,654.0	—	258.0	—
63	1,719.7	38,479.0	(6,740.2)	—
64	1,869.8	—	(10,130.6)	—
65	3,469.0	118,394.3	(5,169.5)	—
66	3,322.1	—	(508.8)	(1,913.7)
67	3,004.3	23,889.1	—	—
68	2,441.4	106,941.9	(12,070.0)	9,745.1
69	2,574.3	(37,998.3)	(8,031.9)	—
70	2,469.6	123,010.3	(14,904.5)	4,213.7
	37,653.4	260,815.6	(186,098.3)	6,925.6

Total Benefit=81,642.9.

Benefit Cost Ratio=81,642.9/37,653.4=2.168.

fade out or disappear?" Does the benefit last in perpetuity? Does the benefit last only a year or two until the lower rate becomes a commonplace, associated with a quality of service that has been customary? Answers to these questions, of course, will have an important bearing on the period over which we should make allowances for the time value of money. If benefits are negative and rate increases drive customers out of the market, does the negative benefit continue until such customers actually leave the market or does it cease at some earlier point, and if so, how do we determine what this point is? In the present analysis we have assumed that benefits continue throughout the study period.

There are additional difficulties which arise as a result of aggregation. We are

considering all communications customers as one consuming unit and that this unit is identical with the group of persons which pays the costs of regulation. We must recognize that the composition of the entire collection of communications customers changes over time. All benefits do not accrue to all people. This is partially due to the fact that not all communications customers avail themselves of all the various communications services which are offered by the carriers. For example, many domestic telephone customers do not make overseas calls. Many persons living today, particularly those in younger age groups, have never sent a telegram. One of our staff members is acquainted with a well-educated person in his thirties who until a few months ago was unaware that there was such a thing as a telegraphic money order,

On the other hand, it seems that since 1960 at any rate, most rate reductions have occurred on the services that most people use most often and the rate increases have occurred on the services that people use least often. However, discriminatory effects are not taken into account. The interstate service which is used by most customers—message toll telephone—can undergo a relatively small rate increase (in terms of the impact on each individual) or undergo an unjustified delay in a rate decrease, so as to offset favorable rate treatment to smaller but privileged groups of customers. This situation may give rise to special benefits for customers receiving certain types of private line services, and it may be virtually impossible to justify such discriminations on either a cost basis or any other basis. One of the most popular such bases nowadays is called competitive necessity. The idea seems to be that discriminatorily low rates favoring a small class of customers are justified in order to foreclose entry into the industry by any new firm which seeks to provide services to such customers. When such discriminations are allowed, the regulatory authorities have in actuality abandoned their responsibility for control of entry and have delegated this function to the carrier which is discriminating. The basic problem, of course, is to determine the appropriate mix of traditional regulation and competitive forces. However, we should note that under existing or probable industry structures, *workable* competition would be of necessity a highly regulated condition. There is no ground for the assumption that a perfectly or workably competitive structure would arise spontaneously in the communications field upon the removal of the present form of regulation. From the behavioral standpoint competition is the struggle to become a monopolist, the entrepreneurial objective being the reduction of uncer-

tainty to the absolute minimum; this minimum point is closely approached by the Bell System today.

The removal of discriminations by regulatory action may have a "wash" effect on the revenue inflow, with some rates being increased and others decreased. Thus the benefit of such removal is not measured. Recent regulatory alleviation of discriminatory effects causes our estimate of benefits to be understated.

We must also consider the problem of jurisdictional offsets. Even though rates may be reduced on interstate communications, the communications companies may rush to the state commissions and successfully persuade them that rates for intrastate communications should be raised. And success at the state level in getting a higher allowable rate of return may be used as a basis for seeking a higher allowable rate of return at the federal level and vice versa. Thus, there is a continual tug of war between the two jurisdictions as the states seek to pull more of the total cost requirement into the federal area and the federal authorities do what they can to resist. The large carriers seem to dominate in determining the particular formula to be used for separations, and the regulatory authorities, in this connection, often seem rather mystified and relatively powerless to do anything about it.

Some consideration should, of course, be given to nonprice benefits. For example, if there are no rate reductions but continual improvements in the quality of service, the resulting benefits will not be measured by our technique. However, it should be recognized that such a situation may put constraints on the consumer's choice. For example, he may prefer less quality at a lower price. This problem can be assumed away by making some provisions concerning the initial demand forecast. Such forecast, in other words, will be assumed to provide for the optimum array

of consumer choices. Negative benefits flowing from an undesirable deterioration of service will, of course, not be measured.

Indirect benefits often flow from the regulatory effort. These include, among other things, the dissemination of information to the public (including the provision of topics and data which engender written articles and dissertations) and the provision of employment opportunities for economics professors. Whether benefits are positive or negative in these cases is a matter of conjecture. Our biased opinion is that benefits are positive.

Finally we come to problems associated with elasticities of demand. These factors should always be taken into account when estimates of revenue inflows are being made. We should consider own-price elasticities, cross elasticities among various services, and income elasticities. All of them are difficult to estimate and all of them are changing over time. One cannot meaningfully state that the own-price elasticity of a given service has a particular value, for elasticities vary with the particular part of the service one is talking about. Short-haul business calls at midmorning on Wednesdays in urban areas may have vastly different elasticities from long-haul residential calls during the Wednesday evening hours in small towns. At the present time, it may be the case that short-haul daytime business calls are more price inelastic than long-haul residential evening calls. The cross elasticity between various types of message toll telephone service and certain kinds of Wide Area Telephone Service may be rather large. Probably also the cross elasticity between WATS and certain types of private line services is of significant proportions. In addition, both telephone and telegraph message volumes are sensitive to changes in real gross national product. Thus, if zero price and income elasticities

are assumed in estimating revenue inflow effects, the results could be seriously in error. The figures accumulated in the various tables so far take into account the influence of elasticities to a certain extent. Probably the estimates are more accurate since the middle 1960's, when serious attention began to be given to the development of computerized econometric models for purposes of making the estimates.

All in all the results of Table 3 indicate that the communications users are not doing too badly as far as federal regulation is concerned. My considered opinion is that the benefit-cost ratio would rise if more resources were devoted to serious study of applied microeconomics as it relates to the regulation of public utilities in general and communications common carriers in particular.

REFERENCES

- Richard Gabel, "The Early Competitive Era in Telephone Communications, 1893-1920," *Law and Contemporary Problems* (Duke University), Vol. 34, No. 2 (Spring 1969), 340-59.
- FCC Annual Report Form M, Schedule 7, American Telephone and Telegraph Company, 1950 through 1969.
- FCC Annual Report Form O, Schedule 414, Western Union Telegraph Company, 1950 through 1969.
- Statistical Tabulations of the FCC Budget and Fiscal Division, 1950-69.
- FCC Annual Report to Congress, Fiscal Year 1953.
- Initial Decision in FCC Docket 11953, adopted July 23, 1958, and made final July 30, 1958. 25 FCC 535, 639.
- Report of the Committee on Governmental Operations, H. R. No. 1836, 90th Congress, 2nd Session, 1968.
- FCC Annual Report Form R, Schedule 414, filed by international record carriers, 1950 through 1969.
- 1970 data taken from company tariff filings and FCC internal documents.

The Social Costs of Regulating the Railroads

By ANN F. FRIEDLAENDER*
Boston College

It is generally agreed that the regulation of the transportation industry by the Interstate Commerce Commission imposes considerable costs upon society. Misallocation of shipments between modes, excess capacity, excessive costs, and poor investment decisions are often thought to arise from regulatory policies. Specifically value-of-service pricing causes many shipments to go by truck that could go by rail more cheaply. Moreover, the lack of rate competition causes a premium to be placed on service, which leads to the maintenance of excess capacity and excessive costs. Similarly, the ICC's rigid insistence on common carrier obligations causes many carriers to maintain capacity far in excess of that used for normal peak loads. Finally, the ICC's attitude toward rate innovations causes railroads to channel their investments into areas that will yield demonstrable cost savings or inhibits them from innovating at all. Thus regulation inflates the total transport bill to society: first, by causing prices to diverge from costs in an inefficient manner, and second, by causing costs to be higher than they would be if all adjustments in output and capacity were permitted.

While these two sources of social costs are obviously connected, for the purposes of analysis it is useful to separate them into static or pricing costs and dynamic or investment costs. The pricing costs repre-

sent the savings that would accrue to society if all transport shipments were reallocated so that each was made at minimum social cost for existing capacity and if all other production and pricing relationships were adjusted accordingly. The investment costs represent the savings that would accrue to society if the existing and future capital stock in the transportation industry were utilized as efficiently as possible.

Measurement of all the dimensions of these costs for all transport modes is probably impossible, given existing knowledge about the costs and demands of the various modes. However, the ICC has collected sufficient data on the railroads to permit estimation of two major sources of social costs arising from rail regulation: those imposed by value-of-service pricing, and those imposed by the maintenance of excess capacity. This paper is concerned with these specific costs.

I. Pricing Costs

Value-of-service pricing is inefficient, not because it causes rates to diverge from costs *per se*, but because it causes rates to diverge from costs on shipments that are likely to have a fairly elastic demand for rail service. Thus the deadweight loss arising from value-of-service pricing is likely to be considerable. Measurement of the deadweight loss from inefficient pricing is well established, e.g., Harberger [4], and generally expresses losses as a proportion of total revenues. For the rail industry, the deadweight loss from value-of-service pricing can be given by

* Conversations with William J. Duffy and Edward J. Kane were instrumental in structuring the problem. I am indebted to Ataman Aksoy, Winston Chiong, and Charles Courtney for research assistance.

$$(1) \quad L_r = 1/2 \sum_i E_{ir} Z_{ir} \delta_{ir}^2$$

where E_{ir} =elasticity of demand of commodity type i for rail services; Z_{ir} =proportion of total rail revenues derived from commodity i ; δ_{ir} =difference between the actual and the efficient price of commodity type i as a proportion of the actual price. I have elsewhere [2] discussed the difficulties associated with measuring these quantities and estimated that these losses lay between 2.0 percent and 1.79 percent of rail revenues in 1964, depending upon whether marginal cost or average total cost was used as the normative price.

Rail operating revenue was \$11,450.3 million in 1969. If we can assume that the percentage losses remained constant between 1964 and 1969, the deadweight loss from rail pricing in 1969 would have been between \$229 million and \$198 million (depending upon whether marginal cost or average cost is used as the normative price).

However, these costs represent only the costs of inefficient pricing to the users of rail services. Since rail services enter into most commodity costs, distortions in rail prices will lead to distortions in other prices. If we can assume that the only distortion in the pricing structure arises from the railroad industry, we can estimate the general cost to society from inefficient rail pricing as

$$(2) \quad L = 1/2 \sum_j \epsilon_j \gamma_j^2 \mu_j$$

where ϵ_j =elasticity of demand for commodity type j ; μ_j =proportion of total value added made by industry j ; γ_j =distortion due to rail pricing. Specifically, $\gamma_j = \rho A_{jt} \delta_{jr}$ where ρ =proportion of rail services in total transport services; A_{jt} =direct and indirect requirements of transport services per dollar output of industry j ; δ_{jr} =divergence between price

and cost for commodity type j in the rail industry.

The 1963 input-output table provides data on A_{jt} and μ_j ; the ICC provides data on δ_{jr} [5]. Unfortunately, data on elasticities are unavailable. However, by assuming a unitary elasticity of demand for all commodities, it is possible to derive rough measures of the general loss arising from value-of-service pricing.

This general loss should be relatively small. For example in 1963 total transport costs (A_{jt}) usually accounted for less than 5% of the costs of any given commodity. Rail services (ρ) accounted for 65% of transport services, while the ratio of the divergence between rail rates and costs to rates (δ) was approximately 0.3 for manufactured goods. If these figures are taken to represent all commodities, the general loss would have been .0044% of GNP in 1963. Detailed calculations of this nature indicate that in 1963 the general loss from value-of-service pricing would have ranged from .0012% to .0044% of GNP, depending upon whether average cost or marginal cost was taken as the normative price. If these losses as a percentage of GNP can be assumed to remain constant, in 1969 the general loss from value-of-service pricing would have ranged from \$12 million to \$41 million.

Thus in 1969 the deadweight loss from value-of-service pricing probably lay between \$220 and \$270 million. While these estimates may be low because of underestimation of the elasticities of demand, a figure of \$500 million doubtless sets an upper bound to the total costs of value-of-service pricing.

II. The Costs of Excess Capacity

There are two ways to assess the costs of excess capacity. The first is to ask by how much costs could be reduced if railroads were permitted to adjust capacity downward to its optimal level for normal pea-

operations. The second is to ask by how much output could increase if existing capacity were fully utilized. These two approaches can be made clear by considering Figure 1. Thus assume that the long-run railroad cost curve is given by LTC and that a given railroad operates along the short-run cost curve STC , which is associated with fixed cost F_a . This cost curve and its associated capacity is optimal for output Q_p in the sense that a railroad operating along STC can produce Q_p at least cost. If, however, actual output is given by Q , the railroad's costs will be given by C , which are greater than they need be. By reducing fixed costs or capacity to F^* the railroad could operate along the short-run cost curve STC^* that would enable it to produce output Q at least cost C^* . The first measure of the costs of excess capacity estimates the difference between C and C^* ; the second measures the difference between Q and Q_p .

In terms of the costs of regulation, it is not clear which is more relevant. While the ICC prevents railroads from abandoning service and divesting capacity, its insistence on value-of-service pricing prevents the railroads from carrying a large number of high-value shipments. Thus in the absence of regulation, it is not clear what combination of divestment and traffic diversion would take place. Therefore, both approaches will be used in assessing the costs of regulating the railroads.

A. Excess Costs

Some time ago Borts [1] pointed out that the relationship between the long-run and the short-run cost elasticities gives some notion of the degree of excess capacity in the rail industry. Specifically, if the short-run elasticity is less than the long-run elasticity, excess capacity must prevail; costs could be lowered by reducing the capital stock. Similarly if the short-run

elasticity is greater than the long-run elasticity, capacity should be expanded to lower costs. Note that the relationship between the two elasticities determines the degree of excess capacity, not the absolute value of either elasticity.

However, the relationship between the short-run and the long-run elasticities only indicates how much excess capacity exists if certain simplifying assumptions are made. Let $C^*=f(X)$ represent the long-run cost curve that represents efficient capacity utilization for any level of output. Let $C=g(X)$ represent the actual short-run cost curve under which a firm operates. E^* represents the long-run cost elasticity while E represents the short-run cost elasticity. Then

$$E/E^* = \frac{dC/dX}{dC^*/dX} \cdot \frac{C^*}{C}$$

and

$$(3) \quad C^*/C = (E/E^*)(mc^*/mc).$$

However, since the short-run and the long-run elasticities and/or marginal costs must be calculated using actual costs and outputs, the product $(E/E^*)(mc^*/mc) = 1$.¹ Nevertheless, the measure can be made operational by assuming that the marginal costs of the two cost curves are equal and attributing the difference in elasticities to the difference in costs. The ratio E/E^* thus becomes equal to C^*/C . In effect this approach shifts the short-run cost curve down (to STC' in Figure 1) until it intersects the estimated long-run cost curve at the output under consideration and measures the difference in average costs between the shifted short-run cost curve (STC') and the estimated short-run

¹ For example consider a linear cost curve $C=a+bX$. Then $E=bX/C$, $E^*=b^*X/C$, $mc=b$, $mc^*=b^*$ and $(E/E^*)(mc^*/mc)=1$. Similarly, consider a log cost curve, $C=\alpha X^\beta$. Then $E=\beta$, $E^*=\beta^*$, $mc=\beta C/X$, $mc^*=\beta^* C^*/X$ and $(E/E^*)(mc^*/mc)=1$.

cost curve (STC). Because it attributes the entire difference in elasticities to difference in average costs, this approach will tend to overestimate the degree of excess costs.

A cross section of 88 Class I railroads was utilized to estimate the long-run cost curve.² The sample period encompassed two four-year averages, 1957-1960 and 1961-1964.³ The railroads were stratified by time (first and second periods) and by region (Eastern, Southern, and Western Territories as defined by the ICC). Covariance analysis was performed to determine where pooling was appropriate. It was found that there was little difference in behavior between time periods, while geographical differences did exist.

For each sample a cost curve was estimated in the form⁴

$$(4) \quad C = a_0 + a_1F + a_2P + a_3F^2 + a_4P^2 + u$$

where F =millions of annual freight-car miles; P =millions of annual passenger-car miles;⁵ C =total costs in millions of 1958 dollars and u =a random disturbance term.

Just as a cross section study should

² However, if most railroads are operating with excess capacity, their actual cost will lie above the optimal costs and a cross section analysis will fail to trace out the minimum costs for any given output. Because of this, the long-run cost elasticities will probably be biased downward, as will be the estimate of excess costs.

³ By using 4-year averages it was hoped that the impact of extreme values could be mitigated and thus avoid the "regression fallacy" that tends to bias cross-section studies. For a full discussion of this problem see Johnston [6].

⁴ Cost curves were also estimated using linear and log cost functions. However the nonlinear cost curves generally performed better in terms of R^2 's, F -statistics, Durbin-Watson statistics, etc. Moreover, only the nonlinear cost curves gave reasonable estimates of excess capacity, which are used in estimating potential output. For a full discussion and the regression equations see Friedlaender [3].

⁵ Freight-car miles and passenger-car miles were used because of the need for data comparability in the cross section and times series analysis.

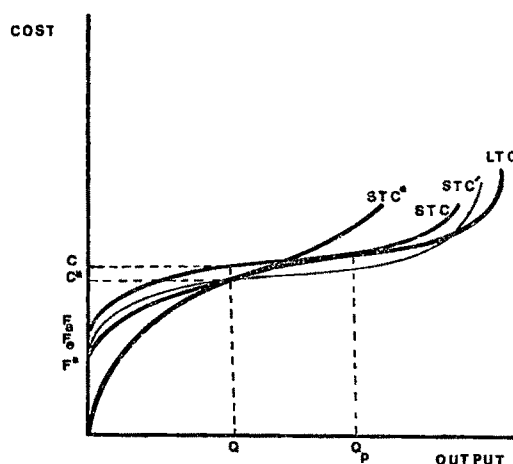


FIGURE 1

trace out the range of minimum costs and thus measure the long-run cost function, a time series analysis performed for any one railroad with a constant capital stock should trace out the short-run cost function. Thus the short-run analysis must be confined to those railroads that exhibited relative constancy of the capital stock during the period under consideration.

Miles of track were used to measure the capital stock of Class I railroads. Of those used in the cross section analysis, thirty-three railroads were found to have relatively constant miles of track.⁶ Short-run cost functions were derived for those railroads.

Quarterly data were used for the period 1957-1964 instead of annual data because of seasonal variations in output that occur over the year. For this reason seasonal slope and intercept dummy variables were added to the equation. Thus, in addition to the form given by equation (4), cost functions for each of the railroads were estimated in the following form:⁷

⁶ Miles of track is a somewhat poor measure of capital stock because it does not reflect quality difference. However it probably gives a fairly good measure of the fixed charges railroads have to bear.

⁷ For a full discussion and the regression equations see Friedlaender [2].

$$\begin{aligned}
 (5) \quad C = & \beta_1 + d_2 + d_3 + d_4 + \beta_5 F + \beta_6 d_2 F \\
 & + \beta_7 d_3 F + \beta_8 d_4 F + \beta_9 P + \beta_{10} d_2 P \\
 & + \beta_{11} d_3 P + \beta_{12} d_4 P + \beta_{13} F^2 \\
 & + \beta_{14} P^2 + v
 \end{aligned}$$

where C =quarterly total costs in millions of 1958 dollars; F =quarterly totals of freight-car miles in millions; P =quarterly totals of passenger-car miles in millions; d_i =seasonal dummy equal to 1 in quarter i , otherwise zero; v =random disturbance term.

Table 1 gives the estimated short-run and long-run cost elasticities and their ratio E/E^* for the thirty-three railroads in the sample.⁸ A direct measure of the costs of regulation can be derived by applying the relationship

$$(6) \quad C_j^* = (C_j)(E_j/E_j^*)$$

for each of the thirty-three railroads in the sample.⁹ Then $\sum C_j^*$ represents the minimum operating cost of the thirty-three railroads while $\sum C_j$ represents the actual cost. Table 1 indicates that for the sample period 1961–1964 the ratio of minimum to actual cost was 0.736. If this ratio can be assumed to hold for all railroads and remain constant over time, the minimum cost of rail operations in 1969 would have been \$6,673.9 million, while the actual cost in 1969 was \$9,066.5 million. Thus, the losses from excess capacity in 1969 were \$2,392.6 million.

⁸ Approximately one-half of the railroads exhibited strong seasonal patterns in their cost relationships and thus had four estimated short-run cost elasticities. Because the ICC prevents seasonal rate variation it encourages highly seasonal use of rail facilities. Therefore, it was felt that the maximum cost elasticities would overestimate the optimal degree of short-run capacity utilization. Consequently, in comparing short-run and long-run cost elasticities, an average of the short-run cost elasticity was used.

⁹ Note that $C_j^* = C_j$; where $(E_j/E_j^*) \geq 1$.

B. Potential Output

Instead of asking by how much costs could be reduced if excess capacity were eliminated, it is useful to ask by how much output could increase if capacity were fully utilized so that the railroads were always operating on the long-run cost curve for their given capital stock. Klein and Preston [7] have developed a useful approach to this question. Thus, assume that it is possible to estimate the following production function

$$(7) \quad Q_a = AL_e^\alpha K_e^\beta u$$

where Q_a =actual rail output; L_e =labor force that would be utilized for efficient production of Q_a ; K_e =capital stock actually utilized; and u =random disturbance term. Given the estimated parameters \hat{A} , $\hat{\alpha}$, $\hat{\beta}$, potential output can be derived as follows

$$(8) \quad Q_p = A\hat{L}_p^\alpha \hat{K}_p^\beta$$

where Q_p =potential output; L_p =labor force required for the efficient utilization of the capital stock, given existing factor returns; and K_p =fully utilized capital stock, which is assumed equal to the existing capital stock K_a .

Rather than calculating the loss in potential output directly as $Q_p - Q_a$, it is useful to estimate the percentage change in output that would arise from the efficient utilization of the capital stock and the labor force. Thus, let $\dot{Q} = Q_p/Q_a - 1$; $\dot{K} = K_p/K_e - 1$; $\dot{L} = L_p/L_e - 1$. Then if the capital stock were fully utilized given the existing labor force, $\dot{Q} = \beta\dot{K}$. However, this does not represent the optimal increase in output, since the output implied by this measure is still off the long-run expansion path. If the combination of labor (L_e) and capital (K_e) that produces actual output (Q_a) represents the least cost combination of factors given relative factor returns, and

TABLE 1—COST ELASTICITIES, ACTUAL COSTS, AND EFFICIENT COSTS:
SELECTED RAILROADS, AVERAGE 1961-1964

Location* and Size	Railroad Name	Cost Elasticities ^b			Operating Costs	
		Short- Run E	Long- Run E*	Ratio E/E*	Actual C	Effi- cient* C*
					(millions of 1958 dollars)	
ES	Bessemer and Lake Erie	.326	.536	.609	10.5	6.4
EM	Boston and Maine	.484	.788	.614	45.8	28.5
EM	Central of New Jersey	1.354	.626	2.163	40.2	40.2
EM	Chicago and Eastern Ill.	.637	1.143	.557	25.5	14.2
EM	Delaware and Hudson	.375	1.047	.358	29.3	10.4
EM	Elgin, Joliet and Eastern	.971	.243	3.984	30.5	30.5
EM	Grand Trunk Western	.811	.799	1.015	41.6	41.6
EM	Lehigh Valley	1.592	.759	2.095	37.1	37.1
EM	Long Island	.281	.731	.384	60.3	23.2
EM	Maine Central	1.463	.692	1.512	16.3	16.2
EM	Pittsburgh and Lake Erie	.514	.297	1.731	25.5	25.5
EL	Baltimore and Ohio	1.209	.899	1.344	264.6	264.6
EL	Chesapeake and Ohio	.838	.910	.920	248.7	225.9
EL	New York Central	1.161	.957	1.213	470.8	470.8
EL	New York, Chicago and St. Louis	.242	.983	.246	85.9	211.4
EL	N.Y., New Haven and Hartford	1.150	.799	1.439	104.2	104.2
EL	Pennsylvania	1.236	.925	1.336	618.4	618.4
EL	Reading	.424	.633	.669	78.1	52.3
SM	Louisiana and Arkansas	1.285	1.080	1.189	14.5	14.5
SL	Atlantic Coast Line	.814	1.169	.695	122.7	85.3
SL	Gulf, Mobile and Ohio	.778	1.009	.771	53.9	41.6
SL	Illinois Central	1.256	1.147	1.199	191.0	191.0
WM	Kansas City Southern	.849	1.572	.539	24.3	13.1
WM	St. Louis and South Western	1.261	1.959	1.154	41.9	41.9
WL	Atchison, Topeka and Santa Fe	.342	2.921	.117	442.7	51.7
WL	Chic., Burlington and Quincy	.765	1.880	.407	188.8	76.8
WL	Chic., Milwaukee, St. Paul and Pac.	.741	1.503	.493	162.4	80.5
WL	Chic., Rock Island and Pacific	1.476	1.747	.845	148.0	125.0
WL	Denver, Rio Grande and Western	.823	1.294	.636	47.3	30.0
WL	Great Northern	1.190	1.599	.744	170.4	126.8
WL	Missouri Pacific	.396	1.733	.228	200.4	45.8
WL	St. Louis and San Francisco	.999	1.475	.667	87.0	58.9
WL	Union Pacific	.727	2.919	.249	338.4	84.3
Total					4466.7	3288.3

(a) E=Eastern Territory; S=railroads with less than 500 miles of track.

S=Southern Territory; M=railroads with 500-3000 miles of track.

W=Western Territory; L=railroads with more than 3000 miles of track.

(b) Elasticities were calculated using average values of cost and outputs for the period 1961-1964.

$$^{(c)} C^* = \begin{cases} (E/E^*)(C) & \text{when } E/E^* < 1 \\ C & \text{when } E/E^* \geq 1 \end{cases}$$

if the production function is homogeneous of degree one, then $K_p/L_p = K_e/L_e$ and $\dot{K} = \dot{L}$. Thus, the output consistent with

the efficient utilization of K_p (given relative factor returns) is given by $\dot{Q} = (\alpha + \beta)\dot{K}$.

To estimate equation (7) it is necessary to derive estimates of L_e and K_e , the labor force and capital stock that represent the efficient production and actual output Q_e . Estimates of efficient capacity can be derived from the estimates of optimal costs. Thus, assume that marginal costs are equal in the short-run and the long-run cost curves. Then the difference between estimated efficient costs (C^*) and estimated short-run variable costs (\hat{vc}) represents the fixed costs consistent with efficient production. In effect, this approach assumes that the short-run cost curve is shifted down until it intersects the long-run cost curve at the actual level of output and measures the fixed costs (F_e) implied by this shifted curve. These are taken to represent the optimal fixed costs as compared to the actual fixed costs F_a (see Figure 1). In the railroad industry capital costs form the bulk of fixed costs. Thus, it is not unreasonable to assume that $F = rK$. Then as long as the returns to capital remain constant, $F_e/F_a = K_e/K_a$, where the F 's represent estimated fixed costs, the K 's represent the capital stocks, the subscript e refers to efficient values and the subscript a refers to actual values. Thus, the ratio between the efficient and the actual capital stock is given by¹⁰

$$(9) \quad \frac{K_e}{K_a} = \frac{C^* - \hat{vc}}{F_a}$$

Having estimated the efficient capital stock, it is then possible to estimate the efficient labor force. With a Cobb-Douglas production function, $dL/L = -(\beta/\alpha)(dK/K)$ for any given movement along an isoquant. The percentage change in capital from the actual to the efficient levels (dK/K) is easily derived from equation (9). However, the only data on the ratio of the capital bill to the wage bill (β/α) is

¹⁰ The variable costs (\hat{vc}) and the fixed costs (\hat{F}_a) associated with the short-run cost function represent quarterly averages.

based on the existing inefficient capital-labor ratios. For the sample and period under consideration, this was 0.662. Since the actual capital-labor ratio is larger than the efficient capital-labor ratio, this figure sets the upper limit of the efficient ratio of the capital bill to the wage bill. For the purposes of this analysis, the ratio β/α was set at two values, 0.5 and 0.33, and the efficient labor force was estimated accordingly.

Thus two production functions in the form of equation (7) were estimated using two different estimates of the labor force. To estimate these equations, the same sample of thirty-three railroads used in the cost analysis was employed. Again, the data was averaged over the sample period to avoid biases due to extreme values.¹¹ Estimates of equation (7) yielded the following relationships:¹²

$$\ln Q_a = .4688 + .1683 \ln K_e$$

(2.906) (3.385)

$$+ .8079 \ln Le_1$$

(15.953)

$$R^2 = .9746 \quad F = 574.518$$

$$\ln Q_a = .5633 + .1306 \ln K_e$$

(3.395) (2.471)

$$+ .8447 \ln Le_2$$

(15.594)

$$R^2 = .9735 \quad F = 551.014$$

where Q_a =railroad operating revenues in millions of 1958 dollars; Le_1 , Le_2 =compensation of efficient labor force in millions of 1958 dollars, where respectively $\beta/\alpha=0.5$ and 0.33; K_e =utilized investment in railroad property used in transportation service in 1958 dollars.

¹¹ The sample only went through 1963 because data on the wage bill by railroad was not available after that.

¹² The figures in parentheses represent t -statistics.

TABLE 2—ESTIMATED FIXED COSTS, ACTUAL AND UTILIZED CAPITAL STOCK
SELECTED RAILROADS, AVERAGE 1961-1963

Location ^a and Size	Railroad Name	Estimated Fixed Costs			Value of Capital Stock	
		Efficient <i>Fe</i>	Actual <i>Fa</i>	Ratio ^b <i>Fe/Fa</i>	Actual <i>Ka</i> (millions of 1958 \$)	Efficient ^c <i>Ke</i>
ES	Bessemer and Lake Erie	.385	2.679	.143	101.9	14.5
EM	Boston and Maine	14.639	19.061	.768	259.9	199.6
EM	Central of New Jersey	17.829	15.356	1.161	179.5	179.5
EM	Chicago and Eastern Ill.	9.231	12.093	.763	115.1	87.8
FM	Delaware and Hudson	40.215	44.965	.894	151.8	135.7
FM	Elgin, Joliet and Eastern	22.949	2.476	9.263	102.9	102.9
EM	Grand Trunk Western	6.136	6.138	.999	125.7	125.5
EM	Lehigh Valley	-2.167	-12.230	5.644	243.3	243.3
EM	Long Island	12.534	20.280	.618	184.6	117.2
EM	Maine Central	-9.117	-11.088	1.216	73.3	73.3
EM	Pittsburgh and Lake Erie	10.483	5.702	1.838	229.8	229.8
EL	Baltimore and Ohio	-67.601	-90.235	1.335	1190.1	1190.1
EL	Chesapeake and Ohio	-59.948	-36.466	.608	1291.4	785.2
EL	New York Central	103.132	75.717	1.362	2286.3	2286.3
EL	New York, Chicago and St. Louis	63.939	81.256	.787	473.5	372.6
EL	N. Y., New Haven and Hartford	1.670	2.459	.769	494.2	335.6
EL	Pennsylvania	260.663	228.054	1.143	3064.2	3064.2
EL	Reading	31.557	35.183	.897	483.5	433.7
SM	Louisiana and Arkansas	7.090	6.304	1.125	69.1	69.1
SL	Atlantic Coast Line	42.725	53.520	.799	576.6	460.7
SL	Gulf, Mobile and Ohio	1.312	4.525	.290	247.6	71.8
SL	Illinois Central	43.121	33.452	1.287	868.3	868.3
SM	Kansas City Southern	-12.163	-9.386	.772	176.3	136.1
WM	St. Louis and South Western	16.387	282.730	.579	216.0	125.1
WL	Atchison, Topeka and Santa Fe	-205.524	-101.817	.495	1998.2	989.0
WL	Chic., Burlington and Quincy	150.044	177.753	.841	957.1	807.8
WL	Chic. Milwaukee, St. Paul and Pac.	-7.363	14.169	.255	895.9	228.5
WL	Chic., Rock Island and Pac.	-121.434	103.957	.856	601.1	514.5
WL	Denver, Rio Grande and Western	8.755	11.999	.730	274.3	200.3
WL	Great Northern	43.282	55.154	.784	881.7	691.3
WL	Missouri Pacific	-59.621	-30.035	.504	969.2	488.5
WL	St. Louis and San Francisco	-103.286	-95.816	.928	442.5	410.6
WL	Union Pacific	10.790	76.602	.140	1628.8	228.0
Total					21858.7	16266.1

^a E = Eastern Territory;

S = Southern Territory;

W = Western Territory;

^b Ratios adjusted to take numerical values of negative numbers into account.^c $K_e = \begin{cases} (Fe/Fa)(Ka) & \text{when } Fe/Fa < 1 \\ Ka & \text{when } Fe/Fa \geq 1 \end{cases}$

S = railroads with less than 500 miles of track.

M = railroads with 500-3000 miles of track.

L = railroads with over 3000 miles of track.

In both regressions $\hat{\alpha} + \hat{\beta} = .97$, although their specific values differ.¹³ In neither case is this figure significantly different

¹³ Note that the ratio $\hat{\beta}/\hat{\alpha}$ cannot be used to estimate the efficient ratio of the capital bill to the wage bill since the specific estimates of $\hat{\alpha}$ and $\hat{\beta}$ depend upon the assumed value of β/α .

from zero. Thus, for the purposes of calculating potential output, it is reasonable to assume that the production function is linearly homogeneous.

Table 2 gives the estimated fixed costs and the actual and efficient capital stock for each of the thirty-three railroads in the

sample. It is thus possible to calculate that $K_a = \sum K_{aj}$, $K_e = \sum K_{ej}$, and $\bar{K} = 34.38$ percent. The maximum output that would employ the existing capital stock efficiently is given by the relationship $\bar{Q} = (\alpha + \beta)\bar{K}$. Thus, the maximum increase in output is 33.35 percent. Since total operating revenue was \$11,450.3 million in 1969, the maximum increase in output would have been \$3,818.5 million. This figure represents the output at which short-run and long-run costs coincide for the existing capital stock. As such it is a counterpart for the estimated excess costs, which attempted to measure by how much costs could be reduced if the capital stock were reduced to permit efficient production at the actual levels of output. This latter figure was estimated to be \$2,392.6 million in 1969.

As pointed out previously, it is not clear which measure is relevant since a relaxation of regulation would doubtless lead to some increase in output and some reduction in capacity. Thus, the true costs of excess capacity probably lie somewhere between \$2.4 and \$3.8 billion.

III. Conclusion

To recapitulate briefly, in 1969 the deadweight loss of value-of-service pricing was approximately \$300 million, while the costs of excess capacity probably lay between \$2.4 and \$3.8 billion. While these estimates are hardly definitive, they do

suggest relative orders of magnitude. Thus, the costs of excess capacity far exceed those of the deadweight loss of value-of-service pricing. In terms of policy prescriptions, this indicates that the pricing distortions caused by value-of-service pricing are considerably less significant than the inability of the railroads to gain additional traffic by pursuing rate competition. In particular, if the ICC relaxed its policy toward abandonments and rate competition, considerable social savings should occur.

REFERENCES

1. Borts, George H., "Increasing Returns in the Railway Industry," *Journal of Political Economy*, Vol. 62 (1954), pp. 316-33.
2. Friedlaender, Ann F., *The Dilemma of Freight Transport Regulation*. Washington: The Brookings Institution, 1969.
3. ———, "Excess Capacity in the Railway Industry," December, 1970 (mimeo).
4. Harberger, Arnold C., "The Measurement of Waste," *American Economic Review*, Vol. 54 (May, 1964), pp. 58-76.
5. I.C.C., "Distribution of the Rail Revenue Contribution by Commodity Groups, 1961." Statement 6-64 (June, 1964).
6. Johnston, J., *Statistical Cost Analysis*. New York: McGraw Hill, 1960.
7. Klein, L. R. and Preston, R. S., "Some New Results in the Measurement of Capacity Utilization," *American Economic Review*, Vol. XVII, No. 1 (March 1967), pp. 34-58.

DISCUSSION

ALFRED E. KAHN: Our chairman has suggested we discuss the papers as a group, in terms of the common problem to which all are addressed—the problem of comparing the cost and benefits of regulation. Actually, the Nelson and Friedlaender papers are so different in approach, I doubt that an extended comparison would be worth the effort.

Dr. Nelson defines benefits wholly in terms of rate reductions realized in a certain time period, and costs solely in terms of the costs of the administrative process itself. In so doing, he is assuming that the efficiency of the communications industry is unaffected by regulation. His sole concern, on the “benefit” side, is with the redistribution of surplus from sellers to buyers.

Professor Friedlaender, in contrast, says nothing about possible benefits or administrative costs, presumably because in the field of transportation both of these are dwarfed by the inefficiencies that regulation has forced on the industry. But neither of the authors addresses himself more than perfunctorily to other important aspects of industrial performance—the quality of the product, the development of new services, the behavior of cost over time, and the degree of discrimination among customers—although Friedlaender does consider these matters in her book. Indeed, Nelson explicitly characterizes all tasks of regulation other than calculation of proper costs and prices as “relatively minor.” Surely most economists would feel, on the contrary, that the influence of regulatory policy on what Leibenstein has called *X*-efficiency is potentially far more important than all its other tasks.

These differences in the approaches of the two papers are certainly reflective, in part, of the very different effects of regulation in the industries they are studying; but there obviously are basic differences in economic perception as well. The basic difference is, of course, that the mere distribution of surplus between sellers and buyers is not a measure of social efficiency, as the economist would define the term, whereas it is precisely social efficiency that Friedlaender attempts to measure.

The only way to determine what kind of difference regulation has made is to compare the actual performance of an industry with one's surmise of what it would have looked like in the absence of government controls. By attributing all changes in rates between 1950 and 1970 to regulation, Nelson in effect assumes rates would otherwise not have changed, which is obviously absurd. Consider the dramatic technological progress and progressive achievement of apparently immense economies of scale in long-distance communications that have, even during a period of inflation, made possible rate reductions which Nelson credits to regulation; and inflation itself, which has required increases in local telephone rates during the same period—increases that, by Nelson's calculus, would all have to be chalked up as demonstrating that consumers would be better off, dollar for dollar, had there been no regulation. This kind of reasoning implicitly assumes, also, that it is the proper function of government to press only for rate decreases, regardless of their impact on the quality and availability of service—a particularly dangerous attitude in time of inflation.

And yet the primitive view is surely not wholly wrong. Regulation is a political institution, and its first political mission is, indeed, to transfer surplus from suppliers to customers. I think the FCC does in fact deserve credit for some large portion of the benefits Nelson attributes to it, and that the result has been an enhancement (though a much smaller one) of social welfare in the economic sense, as well. In my judgment, the historical record in communications amply supports the view that unregulated monopolists cannot reliably be counted upon to reduce rates over time as rapidly as would be justified by cost-reducing innovation, economies of scale and the elasticity of demand. But to assume, without even stopping to argue the point, that long-distance communications rates would have been absolutely constant during the 1950's and 1960's except for the FCC is surely unacceptable.

Professor Friedlaender enjoys an enormous advantage from the fact that the causal im-

pact of regulation on transportation is so clearly observable. My main problem with her paper is that I do not fully understand how much of the second major social cost that she tries to measure is either a total loss or demonstrably attributable to regulation.

Her first estimated social cost is simply the total loss of producer and consumer surplus that results from the excess of prices over marginal cost. I have little trouble attributing this inefficiency primarily to regulation. Her second is what she refers to as the cost of excessive or inadequately utilized rail capacity. I am sure she is right in concluding that these second costs, properly defined, are far greater than the first. As Leibenstein clearly pointed out in his paper on allocative efficiency vs. X-inefficiency, the gain in welfare from a mere reallocation of resources is likely always to be small—the benefit of having a little more spinach and a little less cocoa—whereas, if we use $2X$ resources to do a job that we could do instead with X , the other X is a *real* dead-weight loss. The major waste in transportation is that by holding rates above marginal costs of the various media we are distributing the business inefficiently among them. We are using a lot of $2X$ resources (defined by incremental cost) when a lot of X resources would do the job.

Friedlaender defines these latter inefficiencies as the costs of railroads maintaining excess capacity, which seems to me not the major problem at all. No doubt a more efficient distribution of the transportation business would result in a fuller use of railroad capacity. But the social waste involved in carrying freight at higher rather than lower incremental costs is certainly not the same thing as the cost involved in the railroads being forced by regulation to maintain excess capacity. The proper measure of this latter inefficiency would be the avoidable costs of maintaining that surplus plant. What I need is an explanation of why the avoidable costs are correctly measured, as Friedlaender has done, by the difference between the actual *total* costs of railroad operations in 1969 (not all of which, surely, are avoidable costs) and the estimated minimum cost of doing the same job. Moreover, to the extent that the railroads are ineffi-

cient in doing whatever job they do, I do not think it fair to attribute those inefficiencies solely to regulation—some of them could surely prevail under unregulated oligopoly as well, as Westfield points out. I do concede, however, that if trucking was deregulated and railroads freer to base their rate differentials on marginal costs, much of this inefficiency would be eliminated.

I have trouble also with Friedlaender's alternative measure of the costs of excess railroad capacity—the amount by which sales could be increased if capacity were fully utilized. Since such an increase in sales would entail additional variable costs, surely the potential *gross* increase in output exaggerates the net social loss from idle capacity?

Whether or not these questions about Friedlaender's paper are justified, there remains no doubt that the ICC's resistance to incremental cost rate making has resulted in gross social waste. In the light of this catastrophic expense, one wishes Nelson had been somewhat less cavalier in dismissing the argument of the communications common carriers, particularly AT&T, that they should be permitted to reduce rates down to incremental costs on competitive services. The issue is a difficult one, and complicated by the danger that if AT&T were free to meet competition discriminatorily it might hamper competitive entry and innovation, as Nelson points out. But the horrible example of transportation ought surely to have given him pause before so lightly advocating for communications a policy painfully close to the one pursued by the ICC. In making this observation I should point out that I am a member of AT&T's three-man Economic Advisory Council.

Professor Westfield's paper admirably summarizes my uneasiness with all these attempts at quantitative measurement. But I cannot accept his apparent implication that they are not worth the effort. It seems to me we should grope for a conception of the proper orders of magnitude, in trying to provide guidance for public policy decisions that are going to be made whether we want to think about them or not.

Westfield's paper should instead be taken, I think, as demonstrating how vital it is in each

instance to try explicitly to take into account the limitations of the assumptions underlying the estimates, as I have also tried to suggest in these remarks. It would be an interesting exercise, for example, to try to judge to what extent his warnings truly undermine the results of the Friedlaender paper. I surmise that taking them into account will strengthen, rather than weaken her conclusions. Certainly among the major market imperfections and failures that bear on the validity of her findings are (1) the subsidies to competing transport modes—notably barges and big diesel-burning trucks, (2) the external costs generated by the automobile, and (3) the disproportionate burden of property taxes on the railroads. But all of these make continued value-of-service pricing by the railroads even more socially wasteful than it would otherwise be. This is, I think, the kind of exercise that the Westfield warnings recommend—not abandonment of the effort.

On the other hand, perhaps abandonment *is* the proper prescription when one moves to industries in which competitive organization is simply not feasible. It is surely no coincidence that the estimation efforts Westfield describes have so far been mainly in industries that are believed to be potentially effectively competitive—transportation, where regulation holds price above marginal cost, and natural gas, where it is widely believed to hold it below (see his bibliography). And even in the case of natural gas, I have raised questions elsewhere about the acceptability of the results produced by unregulated field markets. I question the efficacy of bilateral bargaining between regulated oligopsonists on the one hand and producers who are many in number but, nevertheless, organized in various ways for joint or parallel action. The implied case for deregulation must confront also the implication of the fact that gas is both a joint product with oil and directly competitive with oil, whose market is thoroughly cartelized; and it must confront the fact, finally, that deregulation could have substantial effects in redistributing income. All of these considerations are precisely in line with Westfield's observations, and they cast serious doubt on the validity of efforts to estimate the costs of regulation with reference

to a possibly unachievable and socially unacceptable equality of price and marginal cost.

LARRY F. DARBY: I want to emphasize Westfield's assertion that regulation is improperly viewed as an all or nothing proposition. The policy questions should be phrased in terms of the mix of regulatory constraints and free market pressures which maximizes social welfare. Most likely this mix contains a generous dose of regulation, but not necessarily of the present kind. A good case can be made for looking at regulation on the margin. Methodologically it is appealing, but more importantly, it melds well with political realities. The present bundle of constraints was accumulated in piecemeal fashion. Due to the substantial political resistance to regulatory changes, it will probably be dismantled, or amended, in much the same way. Moreover, our limited ability to project the welfare consequences of changes in the bundle requires that we attempt to segregate the effect of specific and marginal adjustments.

My commitment to this view of the issue spares me the acute case of Lipsey-Lancaster fever responsible for Westfield's handwringing pessimism—or realism, depending on your point of view. Ignoring his insistence that we do nothing until the distribution of income is optimal, he argues that in a world of second-best, the marginal product of labor, using the Hotelling approach to estimate the welfare losses from rate-marginal cost inequalities, is zero or negative. The validity of this conclusion depends on the uses to which the estimates are put. Second-best considerations imply that in an absolute sense the estimates will be imperfect surrogates for the "true" welfare effect. But given the quality of the underlying data and the restrictiveness of the underlying assumptions, we already know that. Such estimates should indicate ranges and provide rough starting points for future research and policy proposals. Second-best or not, Friedlaender's fine paper convinces me that capacity, not price, distortions require additional research and regulatory attention. These calculations have important uses if viewed relatively and in proper perspective. Westfield's thoughtful paper should remind us of the need for

being cautious and modest in our claims.

Nelson has a rather eccentric view of the economic and welfare effects of federal regulation of point-to-point communications. After the final adjustment of the data, we are told that regulation is a wise social investment since benefits are more than twice the costs. But, his definitions of the costs and benefits traceable to regulation virtually assure the conclusion. From a social point of view, the final B/C ratio contains a serious upward bias due to a systematic overestimate of benefits and omission of important social costs.

Look at the cost side. Only the "accounting" costs of commission activity and the battery of professionals employed by the companies for adversary purposes are included. But most of us believe that regulatory constraints on the level and structure of rates have more than negligible effects on resource allocation and that some of these effects are properly treated as costs of regulation. The present treatment excludes, for example, such things as the effect of regulation on the input mix, incentives to control costs, incentives to innovate, and incentives to expand into unregulated markets. Now I am not suggesting that Nelson could have resolved these complex issues in this paper, but treating these "hidden" costs as zero biases the outcome in favor of regulation.

On the benefit side, net reductions in carrier revenues overstate benefits, since an unconstrained profit maximizer will ordinarily respond to cost-saving technology by cutting price and increasing output. To the extent that this would have occurred, regulation was redundant.

A more serious problem stems from the fact that with price inelastic demand functions, most of Nelson's benefits are mere income transfers. The social welfare implications of these transfers depend critically on the identity of the gainers and losers. Let us examine the nature of the transfer more closely. Essentially, income is transferred from stockholders of the regulated enterprises and the government (reduced tax revenues) to household and nonhousehold users of regulated services. One can make a good case for assigning negligible welfare implications to the transfer out

of government tax revenues. Furthermore, the transfer from owners to nonhousehold users is not unambiguously a gain in welfare unless the transfer is translated into lower prices and/or better service for households. This leaves the transfer from owners to household users as the dominant source of bona fide improvement in welfare. Assuming (a) that the welfare effects of the transfer from government to users of regulated services is negligible, (b) that the transfer from stockholders of the regulated concerns to nonhousehold users accrues in part to households and in part to other equity holders, and (c) that a dollar transferred from stockholders to households is a dollar gain in welfare, then Nelson's aggregated benefits should be adjusted further. The following expression indicates crudely the percentage of the benefits which qualify under the present assumptions:

$$P = a [t_o (1 - t_p)] + [b (1 - a)] \{a[t_o (1 - t_p)]\}$$

where a is the household users' share of the increase in real income resulting from rate reductions, b is the percent of the increase in real income of nonhousehold users which is passed to households, t_o is the marginal tax rate on corporate income, and t_p is the marginal tax rate on the average taxable income of the owners of the regulated concerns. The first term relates to income transferred directly from owners to households and the second to that transferred via non-household users. Inspection of this expression indicates that P cannot exceed the marginal tax rate on corporate income—that is, $P \leq .5$. Since a , b and t_p are positive and less than one, Nelson's benefits are overstated by a factor greater than two. To work out one simple and generous example, let t_p be .2 and a , b each equal to one. In this case P is .4 and B/C drops to less one. I should point out that P less than one-half is guaranteed by the assumption that a transfer of income from government to household users of regulated services has a negligible effect on social welfare. I can think of no compelling reasons why this should not be acceptable. Even if this group of consumers is found to be especially deserving, regulation of the present variety would

seem to be a relatively inefficient way of effecting the desired transfer.

JAMES R. NELSON: I will start with two unfair and ungracious remarks. The first is that I think discussion of the regulated industries tends to center around the fact of their regulation and not around the special economic characteristics of the industries themselves, or on the economic consequences these characteristics would be likely to have regardless of the regulatory framework in which they operate. This comment is obviously unfair because the topic of this meeting is, precisely: "regulation."

My second remark is that it seems to me that for each speaker separately, and even for all of them together, the assumptions underlying their presentations are so limiting that they eliminate most of the special attributes of the industries they are discussing. But this criticism is obviously ungracious because no one can be expected to read every definition in the dictionary in twenty minutes.

So I must concede that a speaker cannot expect to go from aardvark to zwieback when the rules stipulate one letter only, nor from aardvark to azure when the time available is sufficient to progress only to abacus, for quantitative papers; or to abracadabra, for theory. Therefore what I hope to do today might better be classified as an extension of previous remarks than as criticism.

My first substantive comment is that I am concerned about what appears to be a tendency to import the thought processes and even the methodology of macroeconomics into microeconomic inquiries. Macroeconomics scored its greatest successes, a generation ago, by following Keynes to a very high level of abstraction. But the hallmark of a major leaguer is an ability to hit the curve, and batters armed only with the macro approach have been swinging wildly at the Phillips curve ever since it first appeared. Or, to paraphrase a favorite expression of Paul Samuelson: we must beware of the fallacy of decomposition. The whole does not necessarily break down into the sum of its parts.

The specific application of these generalities to the field of regulation is that microeconom-

ics is not only about *firms* as such—it is also about the *policies* and *operational decisions* of firms. This distinction is especially important in the context of discussions of regulation. A key question, which cannot be answered by approaching the firm as an inert, indivisible, and external object, is this: how, if at all, does the fact of regulation alter the process of firm decision-making? How might changes in the form or objectives of regulation produce further alterations? Specifically, is it possible for regulation ever to achieve an *active* relationship to the regulated industries, in order to affect their efficiency, the quality of their services, etc.? Or must regulation always be *passive*, and if so must it even suffer a certain inevitable economic slippage (over-investment, unduly high cost curves) in the process?

Some of these questions are approached in the papers being considered, but at no point do I obtain any great sense of the distinction between general and specific, or between the firm as such and actual or potential profit centers within the firm. The papers tend to hover on the outer fringes of microeconomics. What may be needed is a plunge into mini-microeconomics.

My second comment is based on a special case of the first. Every paper assumes either one price for the product being sold by the firm (Westfield), or, at least by implication, that prices are the same for all sellers (Friedlaender), or that price structures remain the same when average level is changed (Boyd Nelson).

An exceptionally distinctive feature of the regulated industries is their tendency toward *complicated price structures*, and toward *differential changes* in these price structures in response to differential stresses caused by changes in conditions of supply and demand or in market structure. These complicated price structures have generally stemmed from the same conditions of decreasing cost which originally made competition seem to be an uneconomic or impossible method of social control, and public regulation or ownership the only reasonable alternatives. They have sometimes been imposed (mainly self-imposed) in trucking as part of the argument-by-analogy which is the only kind of economic argument

that can be made for truck regulation. But, whether they proceed from the inside out or the outside in, differentiated rate structures are there; they are probably there to stay; and, in almost every utility, they may often give at least some pale reflection of differential cost structures.

It follows that I can restrain my enthusiasm for the apparent assumption of one price for all units, with the familiar marginal revenue curve at the familiar respectful distance below the demand curve (Westfield), or for the assumption of constant elasticity of demand (Friedlaender), not only in comments like "by assuming a unitary elasticity of demand for all commodities," but in comments like "value-of-service pricing is inefficient not because it causes rates to diverge from costs *per se*, but because it causes rates to diverge from costs on shipments that are likely to have a fairly elastic demand for rail service." Finally, I am disappointed that the assumptions as to demand elasticity, which must be a fundamental ingredient in the estimation of cost savings in the Boyd Nelson paper, are not even mentioned.

To extend the record, if not to set it straight, the following points seem to me to require emphasis. (1) Price structures of regulated industries always have been complicated, and they still are. (2) "Value-of-service" pricing does not just mean pricing at more than marginal cost; it almost invariably means, in practice, pricing at *variable levels or percentages above marginal cost*. (3) Price differentiation and price discrimination need not go hand-in-hand, because the highest rate need not yield the highest unit profit. (4) If the idea of marginal cost pricing is combined with a balanced-budget constraint, which is a principle generally employed in the regulation of profitable utilities and presumably approached in the regulation of less-profitable railways, then "value-of-service" would necessarily imply a complicated rate structure in response to different demand elasticities even in the absence of a complicated cost structure.

My third general comment has already been foreshadowed. It is that regulated industries involve *decreasing costs*. In this type of industry, the distinction between growth and de-

cline of demand can be fundamental. Thus there is an abyss between the Westfield paper, with its emphasis on possible excess *investment*, and the Friedlaender paper, with its emphasis on excess *capacity*. The Boyd Nelson paper brings the issues involved into sharp relief by appearing to show substantial *benefits* from the regulation of telephone service, and *disbenefits* from the regulation of telegraph. One answer to this paradox is surely contained in his passing observation that "many persons living today, particularly those in younger age groups, have never sent a telegram." Among its other problems, then, Western Union must arduously climb back up its cost curve and perhaps, as its main advantage in confronting a host of other problems, AT&T may continue to slide down its cost curve on long distance service.

These industries also involve *costs whose decrease is tied to geography*. (Other things being equal, unit costs decline with increased density of use; but, as officials of Consolidated Edison know to their sorrow, other things may not be equal.) My emphasis on geography may be brought out most clearly by a comment on the Friedlaender paper. First, it is pointed out that "if the short-run [cost-] elasticity is greater than the long-run elasticity, capacity should be expanded to lower costs." Then, in Table 1, we find the following ratios of short-run to long-run cost elasticities, beginning with the highest: Elgin, Joliet and Eastern; Central of New Jersey; Lehigh Valley; Pittsburgh and Lake Erie; Maine Central; New York, New Haven, and Hartford; Baltimore and Ohio; Pennsylvania. If we consult Table 2 for ranking with respect to the similar concept of ratio of efficient to actual capacity, we find the ranking in terms of short-fall of actual below efficient to be as follows: Elgin, Joliet and Eastern; Lehigh Valley; then, after a very substantial gap, Pittsburgh and Lake Erie; New York Central; Baltimore and Ohio; Illinois Central; Maine Central; Central of New Jersey.

The reason for the top position of the Elgin, Joliet and Eastern in this congestion contest is impossible to imagine, but one may still be suspicious of the fact that the Bessemer and Lake Erie, which has the same ownership and

is therefore presumably subject to the same general investment and operating policies, appears in next-to-last place on Table 2 and well below the median in Table 1. The other entries in both lists are in almost every case singularly unlikely candidates for *deficiencies* in capacity. The Jersey Central and the Lehigh Valley, in particular, are completely miscast. The one attribute that the lists seem to have in common is concentration on railroads whose operations are mainly or entirely east of the Mississippi and north of the Ohio. So, whatever the two tables show, it is something endemic to Official Territory. Anyone who studies the history of American railroading since 1920 would guess that that something would not be a deficiency in capacity.

This last sentence leads to a fourth general observation: that, in regulated industries, *place* as an economic factor must be joined by *time*—and by real time, not just by time as abstract chunks labelled “short-run” and “long-run.”

Even at this level of abstraction, it should be noted that for regulated industries the existence of controlled and published tariffs involves the coexistence of something akin to long-run *prices*—there can be little or no relief via the flexibility of spot markets from the implied contract extending via the tariff to every customer and into the indefinite future. This tendency toward arteriosclerosis of the pricing system is greatly obscured by the conventional methods of comparative statics. The most important desideratum may not be a more rational price but a more rational pricing policy.

Emphasis on the time factor may also be brought out by an appeal to the historical

record. It is the fate of many quasi-empirical propositions in economics to gain maximum acceptance only after uncritical followers have failed to notice the original empirical qualifications while the march of events has turned them upside down. So it would appear to be with the Averch-Johnson thesis which is so important to the Westfield paper. A most interesting and scholarly generalization brought forth in a world of 3% interest rates, 6–6½% returns nominally allowed on assets, and actual returns often ranging up from the regulatory “ceiling,” must be approached with extreme caution in an era when the appropriate percentages may read: 8½–10%; 6½–7½%; 5½–8%.

Finally, it may be worth adding a dimension to the Westfield paper with respect to productive factors. Everyone who has worked with the economics of regulated industries is aware that they are highly capital-intensive. But, since their value added tends to be abnormally high relative to value of sales, they may also be abnormally labor-intensive as well. To what extent has this labor-intensiveness interacted, mutually, with both the monopolistic characteristics of many firms in these industries and with their regulated status? Can any such influences be isolated to segregate effects on factor *prices* from effect on factor *costs*? And, when the last passenger train forces off the last Nebraska Congressman at the last isolated station—will it be operated, to this ultimate bitter end, with a full crew aboard? In any case, can a mixture of industry maturity, long union history, regulation, and general political pressure ever be disentangled to give debit where debit is due?

TEACHING ECONOMICS: EXPERIMENTS AND RESULTS

The Lasting Effects of Elementary Economics Courses: Some Preliminary Results

By PHILLIP SAUNDERS
Indiana University

This study found its inspiration in George Stigler's remarks at the 1962 AEA meetings in Pittsburgh, and the results of the studies that G. L. Bach and I were conducting with a nationwide sample of high school social studies teachers at about the same time. In his address, Stigler offered the following hypothesis:

I propose the following test: Select an adequate sample of seniors (I would prefer men five years out of college), equally divided between those who have never had a course in economics and those who have had a conventional one-year course. Give them an examination on current economic problems, not on textbook questions. I predict they will not differ in their performance.¹

The Bach-Saunders studies, limited to high school social studies teachers and using the very elementary Test of Economic Understanding (TEU), tended to support Stigler's hypothesis.² But, given the nature of the sample, the particular testing

instrument involved, and the lack of an adequate measure of several important variables (such as student ability and the grades actually received in college economics courses) the Bach-Saunders results could only be suggestive. A more adequate test of Stigler's hypothesis was needed. Accordingly the present study was designed, proposed, and finally funded by the National Science Foundation last spring. In this *preliminary* report I will focus on: the study design; the sampling procedure; the testing instrument used; and a very tentative look at the *preliminary* data that have been tabulated to date. At the outset I might note that there *do* appear to be statistically significant lasting effects from introductory economics courses. Whether the lasting effects are also educationally significant, however, must be interpreted in the context of the entire framework of learning and teaching; and in this area we still have precious few benchmarks on which to base our judgement.

Study Design

The basic plan of this study is to compare, on a cross-sectional basis, the performance of students who have taken a course in introductory economics with the

¹ George J. Stigler, "Elementary Economic Education," *Am. Econ. Rev., Proc.*, 53 (May 1963) p. 657.

² See G. L. Bach and Phillip Saunders, "Economic Education: Aspirations and Achievements," *Am. Econ. Rev.*, 55 (June 1965) p. 350, and *idem*, "Lasting Effects of Economics Courses at Different Types of Institutions," *Am. Econ. Rev.*, 56 (June 1966) p. 508.

performance of similar students who have not taken such a course at three different points in time: (1) immediately after an introductory course in economics (when most students are sophomores), (2) two years after an introductory course in economics (when most students are seniors), (3) five years after students have graduated from college (some seven years after a sophomore economics course might have been taken).

Sampling Procedure

Since it seemed imperative that some crude but uniform measure of student ability be obtained, it was decided in advance to limit the sample to schools using the College Entrance Examination Board's Scholastic Aptitude Test (CEEB-SAT) as part of their admission procedures. Within this framework, cooperation was secured from 25 schools selected at random in a stratified cluster sample design that insured there would be adequate representation from five "type-of-school" categories, and that there would be adequate representation from schools with different "scholastic potential" in their student bodies.

As fate would have it, the student disruptions at many campuses last spring prevented several of the schools from completing the undergraduate phase of the testing program, and some undergraduate retesting will have to be conducted this spring. There has also been a considerable delay in securing complete alumni information from several schools, so the word "preliminary" in the title of this paper deserves considerably more emphasis than I ever imagined it would when I agreed to present this "preliminary" report last summer.

As it stands, the data in this report represent virtually complete alumni returns from eight private liberal arts colleges, four state colleges and universities,

three predominantly teacher training institutions, two "other" universities and one engineering school. More work must be done before the alumni data from the other participating schools can be incorporated into the final analysis, and I will not attempt to summarize the more fragmentary undergraduate returns at this time.

Testing Instrument

Three sets of printed questionnaires are being used in this study. Each questionnaire is divided into two parts, and is designed so that it can be completed in 40 to 45 minutes on the average. Part I in each questionnaire consists of a series of questions concerning the respondent's name, sex, occupation or class standing, major course of study, whether or not they have taken introductory economics, their present interest in economics as a subject, how important they think economics is, whether or not they think a course in economics should be required for college graduation, and a checklist of items designed to reveal their current reading habits. If the respondent indicates that he has taken or is taking an introductory economics course, he is also asked to indicate whether it is on a required or an elective basis, and he is asked to rate the course, comparing it to other college courses he has taken, with respect to difficulty of subject matter, interest of subject matter, quality of text book, quality of instruction, and time actually spent on the course.

The information obtained from PART I of the questionnaire is supplemented by data from school records on the respondents' CEEB-SAT scores, grades actually received in all undergraduate economics courses taken, the length of these courses, the method of instruction (large section, small section, etc.) and the textbooks used in the various introductory courses involved.

Given this important background data,

the heart of the questionnaire is Part II, which is an especially constructed version of the Test of Understanding in College Economics (TUCE)³ There are four "original" forms of the TUCE, and each contains 33 questions designed to cover several specified content areas and three different types of questions designated as recognition and understanding (RU) questions, simple application (SA) questions, and complex application (CA) questions.

The extensive norming data collected on all 132 TUCE questions was used as a major input in selecting the 33 questions to be used in Part II of the lasting-effects questionnaires. This selection was made during the summer of 1969 when the TUCE committee chairman, Rendigs Fels, Arthur L. Welsh of the Joint Council on Economic Education, who was deeply involved in the staff work on the original test, and I worked intensively with a group of 25 college professors of economics who were attending an economic education research institute for Directors of Joint Council Centers for Economic Education financed by the General Electric Corporation.

The "hybrid" TUCE used in Part II of the lasting effects questionnaires followed the content-coverage and type-of-question specifications outlined by the original test committee. A deliberate attempt was made, however, to omit questions that seemed to rely too much on detailed technical analysis, since it did not seem reasonable to expect alumni who had not taken any economics (or even alumni who had taken an introductory course several years previously) to perform well on purely technical questions.

³ See Rendigs Fels, "A New 'Test of Understanding in College Economics'," *Am. Econ. Rev., Proc.*, 57 (May 1967) pp. 660-666, and *Test of Understanding in College Economics: Manual* (New York: The Psychological Corporation, 1968).

This is an important consideration, since Stigler, when he originally stated his hypothesis, emphasized that the type of question to be used should focus on "current economic problems, not on textbook questions," and he suggested as an example:

Give the student a summary page or two of the arguments and evidence presented in the discussion (in Congress and the public press) of H.R. 5983 (orderly turkeys) and let him explain benefits and costs of the scheme—with the grading based, of course, on the coherence of his argument and relevance of his evidence, not on the conclusions reached.⁴

Unfortunately, the use of this type of essay question in the large scale study design of the project under consideration is administratively and financially impractical. Therefore, we can offer a practical test of only a modified version of Stigler's original hypothesis. Fortunately, Stigler himself was a member of the original TUCE committee and thus had a hand in formulating the test questions we are using. Further, each member of the original committee was asked to look over and comment on the appropriateness of the "hybrid" TUCE *for the purposes of this study*. In this connection Stigler stated:

For reasons that I expressed at some length earlier in the game, I think that the questions are too text-bookish to be a completely honest test of the efficacy of the elementary economics course. However, I see no great harm in a biased examination because I predict that even the maximum estimate of the surviving economic knowledge of an ex-student will be very small.⁵

A set of the detailed specifications for

⁴ George Stigler, *op cit*, p. 659.

⁵ Letter from George Stigler to Rendigs Fels, dated February 13, 1970. This letter was passed on to me, along with the comments of the other TUCE committee members, by Professor Fels, who solicited their opinions on the appropriateness of the "hybrid" TUCE for the purposes of this study.

the test used in this project and the questions themselves can be examined by those with a specific interest in the appropriateness of the test questions in Part II of our questionnaires. It should be noted, however, the simple application questions and the complex application questions of the "hybrid" TUCE, in particular, require students to *use* economics, and even the recognition and understanding questions are not necessarily answerable by rote memory of technical definitions. Further, most of the complex application questions are prefaced by actual or hypothetical newspaper quotations, although they are briefer than the ones alluded to by Stigler above.

Preliminary Results

The 18 schools involved in this preliminary report of the alumni aspects of our study provided a total of 4,790 randomly selected names to which we could actually deliver our questionnaires. As of November 15, 1970, we had received back 1,240 or 26% of these questionnaires. For a variety of reasons, some 181 of these returns are not completely usable. So, at this time, we are left with the results of 1,059 usable alumni returns and some interesting comments written in on some of the unusable ones. Some of these written comments, for example, indicate that both the New Left and the Old Right have successfully recruited at least as far back as the class of 1965.

One female sociology teacher must have spent at least two hours rewording the questions and adding alternatives such as "smash capitalism," "free Bobby," "free Erica," etc., before concluding with this note:

Dear Exploited Research Assistant or Graduate Student who is stuck with coding this mess—(I'm sure P. Saunders himself isn't doing the shitwork).

The terminology and assumptions of these questions are outrageous—and it shows how economics as generally taught in the U.S. never really raises the big questions but merely examines "technical" questions within the framework of the assumptions and values of our "free enterprise" system.

Thank God some of us have met economists in graduate school who helped to strip away the myths we learned in undergraduate economics. You might ask why you are being paid such a lousy salary for the work you are doing now—better yet, you could organize with others similarly exploited to demand pay equal to that of P. Saunders.

One career army officer also spent a lot of time rewriting questions, and he added a condition to one question that stated "if [the government] cannot get the money voluntarily, they have no right forcefully taking the earnings of others, through taxes, to finance their schemes no matter how desirable." Then he concluded:

This slanted questionnaire is representative of the trash taught in the economics courses I took. All study and questions were about which "control" or "tax" should be changed to make public policy—not once did a free market laissez-faire capitalist economy get equal time with the slave systems. No country has gone down the socialist road and survived. Now there is talk of wage and price controls in the U.S. What is next? How can free and non-free economic systems be compared so horribly out of context, and far from reality that they are thought of as being merely forms of resource allocation and not systems allowing freedom or slavery.

Several of the people who sent in usable questionnaires also penned in comments, but not all were as vitriolic as those quoted above. Indeed, a few were almost fan mail, and at least half a dozen mentioned by name outstanding economics professors who had made their introduction to economics exciting and interesting.

TABLE 1—SIMPLE COMPARISON OF ALUMNI SCORES ON "HYBRID" TUCE,
BY AMOUNT OF UNDERGRADUATE ECONOMICS

Preliminary Results

Type of Questions	334 with no econ. (mean SAT score 1155)		399 with only intro. econ. (mean SAT score 1179)		307 with intro. plus other econ. (mean SAT score 1184)		19 with no intro. but other econ. (mean SAT score 1184)	
	mean	(std. dev.)	mean	(std. dev.)	mean	(std. dev.)	mean	(std. dev.)
11 R U	5.1	(1.7)	6.2	(1.9)	7.5	(2.0)	7.3	(1.7)
11 S A	6.2	(1.9)	7.4	(1.9)	8.4	(1.6)	8.6	(1.3)
11 C A	6.6	(2.4)	8.0	(2.2)	9.0	(1.6)	8.7	(1.2)
Total Test	17.9	(4.7)	21.6	(4.9)	24.9	(4.1)	24.6	(3.3)

Test Results

Turning to the 1,059 complete and usable returns, the questionnaires break down into 399 respondents with introductory economics only, 334 with no economics courses at all, 307 with introductory economics plus other undergraduate courses in economics, and 19 respondents who indicated that they had taken no introductory economics, but that they had taken other undergraduate courses in economics. Compared to the overall design of the study, private liberal arts colleges are overrepresented in these preliminary returns. There may also be a slight overrepresentation of schools from the upper end of the "selectivity" scale in these preliminary results.

A simple tabular comparison of the test performance of the alumni responses, broken down by the amount of undergraduate economics taken, is shown in Table 1. A detailed breakdown of these four basic groups on each of the major nontest variables being used in the study has not been performed yet, but in terms of their total CEEB-SAT scores, there are no significant differences in the four main groups shown in Table 1.

There are many interesting comparisons in Table 1 that time prohibits me from commenting on at this point. But, if

we compare the total test scores in column 1 and column 2, there is a difference of 3.7 points in average scores in favor of the alumni with only an introductory economics course compared to the alumni with no college economics. Using a simple t-test, this difference is statistically significant well beyond the .001 level, and this tends to refute the first version of Stigler's hypothesis. Considering the second version, however, one could still argue that 3.7 points on a 33 item test is a "small" difference. Other matter must be cleared out of the way, however, before we can turn to this point. First, the difference might be due to variables other than the amount of economics and average SAT scores shown in Table 1. Second, there is the problem that these responses represent only 22% of the total alumni sample.

In order to handle the "other variables" problem, a multiple regression was run with the information on all of the background variables mentioned above (including sex, type of school, SAT scores, economics beyond the introductory level, etc.) held constant. The regression yielded an R^2 of .498, and it indicated that, other variables held constant, taking only an introductory economics course had a regression coefficient of 3.3 and a t -statistic of 10.8 (which is statistically sig-

nificant well beyond the .001 level.)

It is true that our demonstration of a significant difference in test scores has been established with responses from only 22% of the alumni names provided by the 18 schools covered in this preliminary report. One wishes we had a 100% response, and more attention will be devoted to the problem of nonrespondents later in the study. But, note this: whatever self-selection is built into responding vs. nonresponding applies to *both* those who took an introductory economics course and those who did not. As far as the comparison cited above goes, responding vs. nonresponding is held constant.

Now, is a difference of three or four points "big" or "small"? This is a point on which reasonable men can disagree. The norming data collected on the original versions of the TUCE indicated that introductory economics students who took the test on a "before" and "after" basis improved their scores by about six points. On this basis, a three point difference between those with an introductory course and those without an introductory course indicates that, after some seven years, about 50% of the original gain has been retained. To me this seems like a great deal of retention—at least it seems to be more than most people would have predicted in advance.

In the preliminary regression analyses that have been run to date, the only variables that so far seem to approach the power of the introductory economics variable in explaining differences in total test scores are the CEEB-SAT scores. Using multiple regression analysis to hold other variables constant, each 100 point difference in verbal scores is significantly associated with a difference of 1.12 points in test scores, and each 100 point difference in math scores is significantly associated with a difference of 1.03 points in test scores. On this basis, it would take a dif-

ference of over 300 points in total SAT scores to approach the difference in scores associated with taking introductory economics vs. not taking introductory economics, and I think most people would agree that 300 point differences on these tests are "big" differences.

Introductory Economics Course Ratings

As indicated above, in addition to completing the test portion of the questionnaire, those alumni who had taken an introductory economics course were asked to rate this course, compared to other college courses they had taken, on a series of five variables:

1. *Difficulty of subject matter* (much more difficult, more difficult, about the same in difficulty, less difficult, much less difficult.)
2. *Interest of subject matter* (much more interesting, more interesting, about the same in interest, less interesting, much less interesting.)
3. *Quality of textbook* (one of the very best, above average, average, below average, one of the very worst.)
4. *Quality of instruction* (one of the very best, above average, average, below average, one of the very worst.)
5. *Time you actually spent on the course* (much more time, more time, about the same amount of time, less time, much less time.)

For each variable, each option was followed by a blank, and the respondents' check marks were coded on a scale where 5 was given to the highest option and 1 was given to the lowest option for each variable. The mean values for the respondents' ratings are shown in Table 2.

Since we have no comparable control group with which to compare the ratings in Table 2, the problem of the self-selection of the respondents may be more serious here than it was in comparing test scores above. But, if there is a "halo" effect, it did not wash out differences in the ratings of the various items. I suspect

TABLE 2—SIMPLE COMPARISON OF ALUMNI RATINGS OF INTRODUCTORY ECONOMICS COURSES, BY AMOUNT OF UNDERGRADUATE ECONOMICS

Preliminary Results

Question Compared to other college courses you took, how would you rate your college introductory economics course on each of the following items? (scale used in coding responses: 5=high, 4=above average, 3=average, 2=below average, 1=low.)	399 with only intro. econ.		307 with intro. plus other econ.		Both groups (n=706)	
	mean	(std. dev.)	mean	(std. dev.)	mean	(std. dev.)
Difficulty of Subject Matter	2.99	(0.91)	2.91	(0.78)	2.96	(0.86)
Interest of Subject Matter	2.72	(0.99)	3.40	(0.99)	3.02	(1.05)
Quality of Text Book	3.27	(0.99)	3.57	(0.87)	3.41	(0.93)
Quality of Instruction	3.00	(1.07)	3.31	(1.04)	3.13	(1.07)
Time You Actually Spent On the Course	2.64	(0.91)	2.93	(0.81)	2.77	(0.88)
Mean Rating for All 5 items	2.92		3.22		3.06	

that many of us will be surprised to see that the quality of the textbooks was rated higher than the quality of instruction by both those who took only introductory economics and those who went on for further courses beyond the introductory level. Compared to a mean rating of 3.06 on all five items by both groups, the ratings are relatively higher on quality of text books and quality of instruction than they are on interest in the subject matter and time spent on the introductory economics course. I for one was surprised that both groups in Table 2 rated the introductory economics course slightly below the "average" value of 3.00 in difficulty of subject matter.

Needless to say, no one should bet his life on any of the preliminary findings dis-

cussed in this paper. I am already hopelessly in debt to the many people at the cooperating schools and elsewhere who have contributed many things in many ways to the formulation and execution of the study up to this point. If any of you have any comments, criticisms, or suggestions to make while the study is still in its ongoing stages, they are more than welcome. One of the major purposes of the project is to try to establish certain benchmarks for measuring course improvement in economics. I hope many of you, in addition to critically reviewing the present study, will be moved to experiment with your own introductory courses and see if you can outperform the national benchmarks in the important area of lasting effects.

Teaching Economics to Black Students

By R. GRANN LLOYD
Tennessee State University

Under the dual system of education that developed in the American southland after the Civil War, a system of "Negro Education" evolved. This segregated duality existed at all levels, in both public and private education, from the elementary schools through higher education institutions. It was a system within a system. Since *de jure* segregation was primarily a southern phenomenon, "Negro education" was largely a regional stragem. It was significant, however, since the great majority of America's Negroes and/or black inhabitants lived in the South. And, as Gunnar Myrdal has suggested, the white people of the South "have merely kept Negro [black] education poor and bad."

From the beginning of the dual system of education in the United States until the United States Supreme Court's decision in *Brown vs. Board of Education* (the Kansas case) on May 17, 1954 declared that, "In the field of public education the doctrine of 'separate but equal' has no place," and, to a lesser extent, even today keeping "[black] education poor and bad" has been buttressed by a tight system of racial segregation which has (a) deprived black people of those experiences that cumulatively develop rich cultural backgrounds; (b) denied black people those occupational opportunities and associational outlets that enrich home life and ennoble the quality of life itself; (c) misdirected the work habits of black people generally; and (d) distorted the school performance of black enrollees by providing them with substandard facilities, hand-me-down furniture and equipment, second-hand books abandoned by "white" schools, out-

moded curricula, and by white people giving their blessings to mediocrity in black education. Consequently, the historic black educational experience seems to have been characterized by low achievement in the schools and on standardized tests, where results are greatly influenced by cultural background, work habits, home background, and school performance.

Especially during the last decade black Americans have greatly intensified their long struggle for first-class citizenship in the land of their birth. As a result, millions of Americans have become sensitive to the historic collective neglect and mistreatment of their black fellow-citizens. Many schools, colleges, and universities are increasingly concerned with the black experience and responding to demands for black studies. More significantly, equally as many educational institutions are extending greatly increased opportunities to black students to earn a quality education. This has given rise to much concern about how to teach various subjects, including economics, to black students effectively.

Three Fundamental Questions

It may be profitable to approach the matter of teaching economics to black students by raising three fundamental questions, to wit:

1. What, if any, are the basic differences in teaching economics effectively to black students, on the one hand, and to nonblack students, on the other?

It appears that there are no basic differ-

ences in the teaching of economics to black students and to nonblack students. The subject matter of the discipline is colorless and independent of racial classification. Since the subject matter of economics must be the same for all groups, it would be neither conceptually sound nor practically workable to attempt to teach the discipline differently to black and nonblack students. Even more serious is the possibility that an attempt to teach economics to black students in some manner that is different from the way it is taught to nonblack students may become the cynical gesture by which black brains and effort are diverted from the indispensable exposure to, and involvement in, the mainstream of society where true economic power and financial reward are to be found. Without an understanding of, and the ability to operate in, the societal mainstream, what chance would black people have as a derivative therefrom? Isn't it this superior ability which has been responsible for the success of blacks and whites alike—Robert S. Abbott, Morris Herndon, John Johnson, John Maynard Keynes, Lorimer Milton, Charles Schwab and others? On the other hand, was it not Marcus Garvey and Madame C. J. Walker's efforts to compartmentalize the parameter—thereby limiting the options in financing, markets, and recruitment—that caused their downfall? Nevertheless, a sensitive appreciation of these general matters and the historic milieu which has spawned today's black students suggest that certain factors pose specific barriers to the effective teaching of economics to black students.

2. What factors, if any, pose specific barriers to the effective teaching of economics to black students?

Since it seems unnecessary to belabor the obvious, mention of only three specific barriers to the effective teaching of eco-

nomics to black students should suffice.

First, the dual system of education, which still exists for all practical purposes throughout the United States, has failed to adequately prepare most black students with the basic skills in reading, writing, and mathematics. This relative unpreparedness appears to be, at least partially, responsible for the assumption of a condescending attitude by many instructors—both black and white—toward black students' ability and their failure to offer them rigorous, analytical, mathematical presentations of economic subject matter.

Second, because of their limited exposure to, and contact with, the economic institutions in the community, the limitations of the curricula of the high schools from which they graduate, and the general nature of the environments from which they come, most black students usually bring a smaller inventory of basic economic knowledge into the classroom than middle class white students. This may make economic analysis and the development of economic concepts somewhat more difficult.

Third, far too many black students are not as motivated and/or aggressive as they should be. They raise relatively few fundamental questions in the classroom and seldom insist on rigorous, analytical examination of economic data and/or issues. It appears that perhaps most of them are content to *take* rather than help *make* their education.

The foregoing presentation suggests that whatever specific barriers to the effective teaching of economics to black students exist are probably due to white racism. It also suggests that these barriers may be more quickly overcome if economics instructors would regard the minds of black students as fires to be ignited rather than receptacles to be filled. Indeed, black students, like nonblack students, do not need to be "taught" everything they ought

to know about economics. They should be placed under impelling obligation to read, analyze, and evaluate the basic and the best literature in the field. It suggests further that these barriers will crumble faster if economics instructors will create, by their personal examples, an atmosphere of excellence. The effective teaching of economics, or any other subject, ought to encourage black students to recognize, to desire, and to pursue academic excellence.

Whereas there are no insurmountable barriers to the effective teaching of economics to black students, some of them have been culturally deprived and may need remediation. Those in need of remediation should have an opportunity to obtain it. However, the quality of economics offerings should not be lessened under any circumstances. To do so would have at least two detrimental effects on black students. First, the black students will not be as well prepared as the nonblack students with whom they will be competing in graduate school or the job market. Second, giving black students second-rate education relegates them to a second-class status which is especially detrimental to both their employment opportunities and their self-image.

3. What approaches and/or methods have been found most effective in teaching economics to black students?

Jerome Bruner has raised a basic question which all teachers must face, especially those who teach economics to black students. He asserts, "Students, perforce, have a limited exposure to the materials they are to learn. How can this exposure be made to count in their thinking for the rest of their lives?" Then he says that the answer, "lies in giving students an understanding of the fundamental structure of whatever subjects we choose to teach. This is a minimum requirement for using

knowledge." He continues by arguing that, "in order for a person to be able to recognize the applicability or inapplicability of an idea to a . . . situation and to broaden his learning thereby, he must have clearly in mind the general nature of the phenomenon with which he is dealing."

And yet, to perhaps most black students, much of our effort to teach economics has been characterized by irrelevancy and aimlessness. Too often, in the past, they have viewed the teaching of economics as a disorderly treatment of factual material with no real effort to relate economics to the realities of life in the black community or to reach practical conclusions regarding how a knowledge of economics may be used to improve the well-being of its inhabitants. This "aimless accumulation of precise knowledge, inert and unutilized," as Alfred North Whitehead characterized it, has often "turned off" the interest of black students and probably led to some paralysis of economic thought by many of them.

If economics is indeed an art and a science then it should be made applicable to all situations and conditions and not confined to any spectrum of color or bias. Therefore, there should be no fundamental difference in the way economics is taught to black students and the way it is taught to nonblack students. At the same time, however, some approaches to the teaching of economics to black students appear to be more generally effective than others.

Economics instructors in black colleges and universities seem to be making more and more use of multi-media representation as they attempt to increase black students' appreciation of economics, to create impressions that stimulate their interest in the discipline, and to facilitate their comprehension of the subject. Thoughtful and well-planned use of multi-media represen-

tation is being effectively used in bridging the gap between what seems to be the remoteness of what the students are studying in economics and the world of reality as black students perceive it. Furthermore, the use of various types of multi-media representation with black students not only enhances economics teaching and learning, but may facilitate the economical use of time in the teaching-learning process.

Although multi-media representation is not strictly the same as the reality, it is being used more and more to satisfy the need for greater realism in teaching economics to black students. In general, it appears that an ever-increasing number of instructors of economics in black colleges and universities are using multi-media representation in teaching situations involving motivation, reality, and efficiency. The medium is also useful in creating general impressions of economic situations and conditions. It is especially helpful in enhancing the understanding of economics by those black students, whose experiences with the enriching aspects of our society's cultural and economic institutions have been limited, since it enables them to see and/or hear "the real thing" or a dynamic representation thereof.

Some instructors in black colleges and universities seem to be finding that programmed instruction in economics is quite effective with black students. In programmed instruction (a) instructional objectives are clearly specified, (b) subject-matter is carefully sequenced, (c) subject-matter is usually presented in small incremental steps, (d) the active participation of the learner is required, and (e) immediate feedback regarding the adequacy or correctness of response, is provided. Hence, this approach should satisfy those who contend—rightly or wrongly—that the effective teaching of economics to black students requires that course mate-

rials be highly organized. Nevertheless, the ordering of course materials in logical sequence and giving careful attention to course structure do not necessarily result in a given course being so highly structured and/or academically oriented that intellectual initiative is stifled.

The teaching of economics to black students with programmed materials seems to be effective for at least three reasons, to wit: (1) programmed instruction recognizes individual differences and allows the learner to proceed at his own rate; (2) the discipline within programmed instruction is conducive to learning—it accentuates the organized nature of knowledge and forces the learner to actively participate in his own learning; and (3) programmed instruction removes anxiety over possible failure and provides instant awareness of results. These things are especially important to most black students, toward whose ability to learn economics their instructors have often taken a condescending attitude.

The inquiry-centered approach to the teaching of economics to black students seems also to be effective. With this approach students and instructors have cooperatively selected persistent economic issues, policies, and institutions for examination and identified economic concepts, principles, and relationships that may clarify and/or explain them. These economic concepts, issues, and relationships then serve as organizing principles that determine the selection of course content. Thus the basic principles and phenomena involved in clarifying and explaining the matters in this area that are of special interest to black students provide a setting which is conducive to their understanding of economic principles, economic institutions, and economic policies and processes. As black students inquire into these matters they discover the processes by which economic ideas and policies are

developed, verified, calibrated, evaluated, and refined.

The inquiry-centered approach forces students to (a) define issues and important concepts, (b) develop inquiry strategies for attacking issues and problems, (c) collect authoritative data for use in clarifying and solving problems, (d) develop criteria for acceptance or rejection of an original hypothesis, and (e) reach conclusions based on thorough examination of available data.

By capitalizing on black students' interest in issues and problems of concern to the black ghetto, the inquiry-centered approach may be used to force black students to actively participate in their own education; to emphasize how economic data and analysis may be used to clarify issues and to develop valid explanations of economic behavior, institutions, and policies; and to facilitate black students' understanding of economics.

On the basis of observation and report, there seems to be a considerable amount of field work taking place in black colleges and universities as instructors seek to make economics more meaningful to black students. The number of in-the-field, problem-solving courses where economic analysis can be applied to issues and problems, such as (1) housing in the black ghetto, (2) the transportation dilemmas of the residents of the black neighborhood, (3) the market in the black community, (4) barriers to effective manpower utilization, (5) the economic effects of white racism on selected institutions in the black neighborhood and ways of combating them, and (6) the impact of the prevailing local tax structure on the black community and so forth seem to be increasing in black educational institutions.

Perhaps the most meaningful approach to the teaching of economics to black students, or to any students, is to make the subject matter relevant. This approach re-

quires that economics be taught from a cultural context in which the focus for considering the subject matter of the field, including the basic principles of economics, directly involve the black experience. This approach involves showing black students how the rigorous application of economic theory and concepts may be used to analyze and solve problems in the black community, and demonstrating how the basic areas of economics are of importance to black students.

The object of cultural context teaching of economics is not to constrict students' learning to the boundaries of the black experience, but to utilize their interest in that experience as a motivational leverage for the teaching and learning of the discipline. For example, in black colleges and universities (a) the study of inflation is often centered around its impact on the black ghetto; (b) the concepts of "leakage" and the "multiplier" are applied to the analysis of increased investment spending in the black community; (c) emphasis on the economics of discrimination is a characteristic of the labor economics course; (d) poverty is emphasized in the course on contemporary economic problems in the United States; (e) the course on the economic history of the United States concentrates on the economics of slavery; (f) the course on economic development emphasizes the problems of developing an economic foundation for black people in the United States.

More effort is probably put into the teaching of economics in black colleges and universities than appears to be true of the larger institutions. Much more attention is probably given to the selection of examples to which black students can relate. For instance, in the study of tariffs, segregation is often analyzed as a tariff; little stores in ghetto neighborhoods are examined in connection with the study of imperfect competition; interest in the

multiplier principle is frequently aroused through discussion of community multipliers and the problems of black banks in retaining reserves; and similar examples.

This approach to the teaching of economics courses to black students does not require lowering the quality of economics courses, or making them less rigorous and theoretical. It does mean that within the limits of the general principles of teaching and learning considerable latitude in adaptation to the realities of society and the needs of students is both possible and desirable. Indeed, this approach is in accord with Charles A. Beard's (slightly paraphrased) assertion that, "instruction in [economics] is conditioned by the spirit and letter of scholarship, by the realities and ideas of the society in which it is carried on, and by the nature and limitations of the learning process at the various grade levels across which it is distributed."

Hence, "the necessities of scholarship, the realities of society, and the requirements of the teaching and learning process" are the stubborn and irreducible stipulations that are binding upon teachers of economics as well as teachers of other subjects. This seems to indicate that economics and/or any other subject can be taught in an intellectually honest manner to any student on any grade level without regard to race, creed, or color. At the same time, however, some economics instructors in black colleges and universities, who have taught both black and white students, suggest that economics course materials for use with black students should probably be more highly organized than materials for use with white students. In this connection, it seems significant to note that John F. Due asserts "there must be a high degree of selectivity in presentation" of the major economic concepts and information about the economic system

without making any reference whatsoever to race.

Summary

The foregoing discussion sets forth in bold relief the position assumed here, to wit:

1. There should be essentially no difference in the way economics is taught to black students and nonblack students. However, because of the deeply ingrained and continuing debilitating effects of 350 years of slavery and second-class citizenship the effective teaching of economics to most black students may require the utilization of carefully selected approaches. Albeit, whatever problems exist in the teaching of economics to black students are marginal problems—reading disabilities, comprehension insufficiency, and mathematical inadequacies, all resulting from inferior educational opportunity plus cultural deprivation, but all susceptible to being overcome by hard work, determination, and good teaching.

2. The constant complaint about black students' mathematics insufficiency is probably not as serious as is generally believed, since mathematics itself is necessary only for the higher levels of economic theory. And, as Paul A. Samuelson has pointed out, "Logical reasoning is the key to success in the matter of basic economic principles, and shrewd weighing of empirical evidence is the key to success in mastery of economic applications."

3. There are significant differences in the backgrounds and motivations of black and nonblack students. This has led to the highly arguable contention by some educators that teaching economics to black students may be effectively approached with somewhat less abstraction and in a more practical manner than with nonblack students.

4. Many, perhaps most, black students

are capable of learning economics as rapidly and as well as any other students, but the lily-white manner in which it has been taught has stifled their ability to demonstrate how really capable they are to learn economics.

5. The underlying premise set forth by Bruner in *The Process of Education* seems to be that "the first object of any act of learning, over and above the pleasure it may give, is that it should serve us in the future. Learning should not only take us somewhere; it should allow us later to go further more easily." Hence, Bruner advocates the transfer of attitudes and principles which "consist of learning initially not a skill but a general idea, which can then be used as a basis for recognizing subsequent problems as special cases of the idea originally mastered." This sort of transfer "is dependent upon mastery of the structure of the subject-matter," which Bruner defines as "understanding it in a way that permits many other things to be related to it meaningfully. To learn structure, in short, is to learn how things are related."

Unfortunately, we seem to have only inadequately, if at all, taken this premise and these guidelines into account in teaching economics to black students. Moreover, the lily-white approach to the teaching of the discipline has deprived most black students of an opportunity "to learn how [economics is] related" to the realities of their lives and the viability of their communities.

6. Whereas the desire for effectiveness seems to suggest that approaches to the teaching of economics to black students should probably be very carefully selected, and perhaps somewhat differenti-

ated from those used with nonblack students, the substance and rigor must be the same. What must never be forgotten in teaching economics to black students, or in teaching any subject to any students, is that:

There is no procedure that can render substance unnecessary; there is no technique of classroom legerdemain that can take the place of scholarly competence; there is no device of instruction that can raise the quality of the educative process above the purpose, the knowledge, the understanding, the vision of the teacher who employs it.

REFERENCES

- American Historical Association, *Report of the Commission on the Social Studies, Conclusions and Recommendations Of The Commission*, New York: Charles Scribner's Sons, 1934, pp. 71-72.
- Charles A. Beard, *A Charter For The Social Sciences*, New York: Charles A. Scribner's Sons, 1932, pp. 2, 3.
- Jerome Bruner, *The Process of Education*, Cambridge, Massachusetts, Harvard University Press, 1961, pp. 11, 17, 18.
- John F. Due, "The Teaching of Economics: Goals, Testing, and Evaluation," *New Developments in the Teaching of Economics*, edited by Keith G. Lumsden, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967, p. 111.
- Gunnar Myrdal, *An American Dilemma: The Negro Problem and Modern Democracy*, Volume 2, New York: Harper and Row Publishers, 1962, p. 896.
- Paul A. Samuelson, *Economics: An Introductory Analysis* (Seventh Edition), New York: McGraw-Hill Book Company, 1967, p. 5.
- Alfred North Whitehead, *The Aims of Education*, New York: New American Library (Mentor Books), 1949, p. 48.

Videotaped Dialogues in Economics*

By BARRY CASTRO
Hostos Community College

The videotaped dialogues on which I am reporting were primarily intended to contribute to the creation and/or enrichment of intellectual interaction by opening new channels of two-way communication among faculty and between faculty and students.¹ They were also intended to stimulate viewer interest in the skills of dialogue, in the grammar of television, in the goals of education, and in the substantive problems with which the dialogue dealt. In each of these areas, the tapes were intended to initiate a process rather than provide specific answers to specific questions. It was hoped that they might eventually provide an overdue antidote to the single "right" answer bias of both programmed instruction and much of our testing.

Taped dialogues do not require the assumption that students are generally interested in our discipline (or any discipline). Their production was premised on several other assumptions:—that problems exist in which students are interested; that it is possible for us to find out

what these problems are and make ourselves and our disciplines relevant to them; that if we do this, disciplinary sophistication will follow because our students will want it, and, that while we can create environments in which such a desire is more likely to develop, we cannot effectively impose that desire on our students. These assumptions embody an educational philosophy with which I believe taped dialogues are strongly identified. However, given both the enormous amount that has been written about general education, liberal education, student-centered education, etc. and the time and space constraints of this paper, I will not belabor that identification here.

Our reliance on dialogue was critical. It was achieved because we chose panelists whom we expected to have strong methodological and/or substantive differences and because we edited in a manner intended to highlight these differences. Dialogue was emphasized because it is compatible with the problem-centered approach I believe to be generally more responsive to student interests than the traditional disciplinary orientation—because it seems inherently more interesting (and consequently more likely to secure student participation without coercion) than most alternative techniques; because it emphasizes the limited relevance of narrowly defined "right" and "wrong" answers to problems; because it makes it clear that faculty members can be guilty of a wide variety of errors and reduces whatever inhibition to learning may be associated with the wide gulf students often feel exists between themselves and faculty; because it makes it more difficult for

* Many people are responsible for whatever success this project has had. However, my report would be incomplete if I did not at least single out the contributions of George Braun, Roger Folsom, R. E. Kayton, Roy Moats and Virginia Nousse of the Naval Postgraduate School, William B. Daniels of Hostos Community College and Kenneth Weingarten of the Human Resources Organization. It seems clear to me that the project would not have happened without them.

¹ I should like you to note that whatever communication enhancement is made possible by the procedures I will recommend should be realizable in a much wider variety of settings than that defined by an academic department, and that the academic uses of these tapes is in no way peculiarly relevant to economics. The fact that our dialogues were dominated by academic economists was solely a result of my own background.

individual faculty members to insulate themselves from the criticism of their colleagues and, consequently, makes it more difficult for them to maintain rigid positions; and finally, because the student's exposure to the norms which make possible a constructive discussion between panelists with varying frames of reference and/or conclusions is expected to make it easier for him to enter into a dialogue of his own.

There were many good reasons to put these dialogues on videotape. However, the one that probably got me started was the enormously insensitive use of the medium —of which academics have generally been guilty. Most academic television has been a simple recording of the traditional lecture on videotape. Too often, when students have let it be known that they were bored by an hour of looking at a picture of a man talking to them, or angered at the attempt to depersonalize education further by eliminating the possibility of two-way communication with their instructors, their responses were considered to reflect the inappropriateness of television for academic use. I cannot stress too strongly that this failure has been a failure to exploit the medium rather than a failure of the medium itself.

On the positive side, videotaped materials facilitate presentation to audiences small enough to begin their own follow-up dialogues immediately after the tape has been presented. The editorial capability of the medium makes it possible for the original dialogues between faculty-panelists to continue over very extended periods without making it necessary for the audience to sit through very extended viewings, and permits re-editing of the same original taped material so as to make it suitable for a wide variety of potential audiences. Its editorial capability also makes possible continued amendment of the original dialogue. If a panelist were to think of the re-

buttal to an argument he should have used but had overlooked, it could be added to the tape. Similarly, his adversaries would be able to insert their rebuttals. The possibility of viewing taped material immediately after it has been recorded gives each panelist an opportunity to see himself in action and respond to whatever he finds wanting in his own presentation while the dialogue is still fresh in his mind. Moreover, television permits repeated showings and multiple simultaneous showings. Finally, it depends on production facilities and presentation equipment which are generally available in American colleges and universities. These advantages are important and we tried to use them fully in our project.

Our procedure was to bring together a group of economists who were mutually interested in some question but were expected to disagree. They met in an informal, somewhat unstructured situation to talk that problem over for an entire weekend. Everything was taped. Occasionally the panelists would look at some of their own recently completed dialogue to get a bit of an outsider's perspective on what needed to be said. If the discussion seemed too diffuse and rambling, one of the panelists with a relatively strong position would be asked to summarize his views for five or ten minutes after which the group would react to them. If the group seemed to be fixed on a relatively minor point or the spontaneity of the discussion seemed to be breaking down, we would ask students who were monitoring the proceedings from a separate room to direct some questions at the panelists. In general, what we ran seemed to us to be a kind of academic encounter group from which we at least came away with the feeling that something meaningful had passed between our colleagues and ourselves. The infrequency with which this happens in ordinary intradepartmental interaction

was often remarked on and seemed an important secondary benefit of the project.

Our panelists were drawn from about a dozen different institutions. Finished tapes were shown to students at most of them. We did this because we wanted to encourage further experimentation with these techniques in a number of schools. However, we would recommend that future projects of this sort attempt to rely on faculty inside a single university and that finished tapes be used only in the schools at which they were produced. While this procedure may sacrifice some expertise and requires a much greater total production effort, it has numerous advantages: (1) The dramatic quality of the dialogues is likely to be enhanced if the participants are known to the viewers. (2) Continuing interaction is fostered if the viewer has relatively easy access to panelists for followup discussions. (3) Continuing interaction is also fostered if the panelists have access to each other after taping has been completed. (4) The development of a new channel of intra-university, inter-disciplinary communication could serve to emphasize broadly intellectual, rather than exclusively disciplinary matters. (5) The contact between participating professors and television staffs opened up during production might easily be continued—hopefully, resulting in additional experimentation with television by individual faculty members.

The procedure we found ourselves moving toward was to begin with a departmental discussion intended to identify those issues which seemed to the department to be (1) of general interest, (2) relevant for classroom presentation, and (3) likely to provoke some disagreement among the participating faculty. One member of that faculty would act as project coordinator and interview each of the other participating faculty members separately on the question at hand. These in-

terviews were expected to produce stronger and more cohesive statements of positions than would have been forthcoming had they been solicited during a group session. Each interview was to be taped and subsequently edited into a single short tape in which areas of agreement and peripheral issues were to be edited out and disagreements in approach or conclusions were to be resequenced in a manner which made them very difficult to overlook. The resulting tape would be shown only to the participating faculty members (meeting as a group) who would be asked to correct any misimpressions which might be due to the editing and subsequently asked to discuss each other's positions as they were stated on the tape they had been shown. That discussion would itself be videotaped and edited into one or several short tapes for use at the school in which they were made. If possible, we suggested that a small group of students (one or two nominated by each participating faculty member) view the tapes before they were used in the classroom and embark on their own discussion of both the substance of what they had seen and its relevance for the classroom. These student sessions would also be taped and material from them could be integrated with material from the faculty tapes for classroom presentation. The emphasis of this procedure is on the communication channels opened up within the university rather than on the finished tape as a final product.

Whichever production scheme is used, I very strongly recommend that your editorial needs be considered in advance and that you utilize the editorial capability of the medium. We need not emulate most current university productions in confining our editing to cleaning up some rough spots or cutting out parts of our tape so as to fit it into available time constraints. Editing can, and in my view, should strive for fast-paced one-liners interspersed with

some longer expositions to enable the viewer to catch his breath. The relative weight you place on a continuous flow of conversation as opposed to a staccato "Laugh-In" rhythm should be based on some experimentation with your audience. My own experience suggests that the first tapes you produce should be relatively conservative and that your audience will be increasingly receptive to tightly edited discontinuous material as they become accustomed to the sophisticated use of the medium in an academic context. I should also add that you are likely to find heliscan (one inch) equipment inadequate as you move toward really tight editing and that quad-head (two-inch) equipment will relieve you of many technical problems if you are fortunate enough to have access to it. Finally I expect that you will find it impossible to undertake efficient editing unless you have arranged for a transcript to be made of everything that was said on camera. We found it most efficient to do our preliminary editing on the basis of the transcript alone.

The editorial effect we found most successful was the conterposition of opposing points of view so that the differences between our panelists were difficult to miss. We also found that we could do away with the constraints imposed by a moderator by simply allowing conversations to flow and letting our editorial capability assume the moderator's role. It was generally possible to use the same original material to produce a variety of tapes intended for different audiences or to focus on somewhat different problems. In no case did we try to follow through an argument to some definitive conclusion. What we sought was the clear presentation of whatever problem was at hand and at least two approaches to its resolution.

We undertook no systematic evaluation of our efforts during this first year of our experimentation with taped dialogues. The recommendations I have made are

based on my own response to our various efforts and to the feedback I and many of our panelists have had when we used the tapes in class. The possibilities of self-deception in so loose a feedback system are clear and there is no question but that systematic evaluation of our recommendations needs to be undertaken. While I am no longer at the Naval Postgraduate School where our dialogues were produced and, consequently, am no longer directly involved with this project, an attempt at such an evaluation is under way and I trust that its results will be made available to you without undue delay.

I should like to point out to you that I am interested in extensions of the basic approach I have outlined here—that William B. Daniels, one of my colleagues at Hostos, and I are applying this approach to the development of communication channels between a school and the community it serves, and that if any of you are interested in that problem, we would be happy to let you know more about what we are doing. We would also be happy to share whatever sophistication we have acquired about handling the detailed production problems of videotaped dialogue with you. Finally, I should like to repeat my request that any of you pursuing similar projects let us know about them. It would be an unpleasant irony if this paper was to be no more than a one-way communication.

We had guessed you might be interested in an example of one of the taped dialogues we made last year and one will be shown to you in a few moments. I should like to thank the panelists (Roger Folsom of the Naval Postgraduate School, Douglas Korty of the Collective of Socialist Economists, John Lindauer of Claremont, Jack Michaelson of the University of California at Santa Cruz and Lorie Tarshis of Stanford) and our sponsor, the Bureau of Naval Personnel, for their permission to show the tape to you.

DISCUSSION

IRVING MORRISSETT: The Saunders Project to test the "Stigler hypothesis" is a heroic effort. Which of us expects any measurable results from our fondest enterprises—immediately or two years later, much less seven years later? Which of us would like to have a test made seven years later of the effect on a student of reading our favorite book—especially if we wrote it ourselves? Which of us would like to have an objective measurement of the results of our efforts as good parents, compared with those of our unconcerned neighbors next door?

But Saunders and his sponsor, the National Science Foundation, are brave or foolhardy, or both. They reflect a trend—beginning in the U. S. Office of Education, the National Science Foundation, and the school districts of the nation, although not yet the universities—that might sweep the country, causing acute discomfort to educators at all levels. It is a trend that challenges the comfortable but unlikely assumption that educational output is proportional to input, that demands "accountability" and performance measurements.

Saunders' investigative methods are sound and his results are interesting and convincing, as far as they go. There is no reason to doubt that additional returns, as they come in, will continue to be interesting and convincing. Whether a 3.7 point gain on a 33-point scale is sufficient encouragement to make us continue to offer the introductory course is open to debate. This figure *is* statistically significant, thereby casting doubt on the Stigler hypothesis, and it is half as much as the retention shown by previous tests given immediately after the introductory course. Note also that the reported 11 percent gain (3.7/33) can be boosted to 15 percent (3.7/25) if the score of eight that would be obtained by purely random guesses is subtracted from the base of 33. The gain factor can be further boosted to 24.5 percent if it is taken as a percentage of the potential gain (3.7/15.1).

Also of interest is the finding that ability of respondents is related to test scores, although weakly. One hundred points, on either the verbal or mathematics SATs, is worth about one

point on the test. Further similar results can be expected as Saunders' research is pursued, throwing light both on the main objective of testing the direct effect of the introductory economics course and also on the effects of control variables such as the SATs, the type of college or university attended, and current occupations of the responding alumni. As is well known, the effects contributed by uncontrolled variables often supply some of the most interesting results of a research effort.

I have always had some doubts about the TUCE (Test of Understanding of College Economics). Thirty-three multiple-choice questions seem like a slender base on which to judge all the substance of two semesters of economics. Subtracting eight points for correct answers that could be gotten by chance, the spread is only twenty-five points; and a few more points need to be subtracted from the base to account for correct answers that almost anyone could get by applying some common sense. There is also the nagging doubt as to whether multiple-choice questions can measure achievement with respect to the analytical skills that one hopes to nurture in the introductory economics course.

Despite these doubts, I have become convinced that the TUCE is a good test. The individual questions are well constructed; they reflect levels of understanding from simple to complex; and they have been subjected to broad criticism and testing in the field. The "hybrid" test used by Saunders was made up from all of the original four TUCE tests and is at least as good as and probably better than any one of the four.

Responses to Professor Saunders' questionnaire have yielded some colorful bonus results. The volleys from the Left, nominating Professor Saunders as capitalist pig professor of the week, make a colorful contribution to the literature. The volleys from the Right, nominating Professor Saunders as the socialist Commie dupe of the week, are equally eloquent. If sufficient similar data are available, it might be possible to test the hypothesis that, in the heat of emotion, Leftists are more likely to employ four-letter words than Rightists. At any rate, after noting the cognitive and affec-

tive impact of these bonus statements from Left and Right, Professor Saunders was no doubt heartened by middle-of-the-road supporters who recalled their experiences in introductory economics courses with pleasure.

Having lavished a number of supportive and approving comments on Saunders' study, it is time to unleash my critical comments. My one and only criticism, and a very important one, concerns the implicit conservatism of the study—conservatism, not with respect to content, but with respect to methods of teaching. Saunders is, of course, measuring the effects of the courses taught, based on the usual content and the usual methods of teaching. The criticisms from the Right and Left as to the content of the introductory course are a little bit off the mark, but not entirely so. It is probably fair to assume that Saunders is not entirely opposed to the content of the usual introductory course; if he were so opposed, one might suppose that he would not waste his time conducting such a research project or, at the least, he would find occasion to express his opposition to the content. Similarly, it can be assumed that Saunders is not completely and passionately opposed to the *methods* being employed to teach introductory economics courses.

My own view is a drastic one. I think that the typical methods of undergraduate teaching in American colleges and universities, including the teaching of introductory economics, are backward, unimaginative, and disgraceful. The best defense that can be given for them is the dubious argument that this is the way that it has been done for a hundred years and more, and the somewhat more cogent argument that this method is much better than no method at all. I believe that teaching methods—or rather the whole teaching-learning configuration—in our entire educational system must and will change drastically in the coming decades. I would like to see college professors at the forefront of this movement rather than bringing up the rear as they now are.

My interest in the Saunders study, therefore, is in finding evidences of effects caused by variations in the teaching-learning configuration. The control variables introduced in the study are not very promising in this respect.

Respondents are asked to indicate whether the course was required or elective, to make a judgment about the quality of instruction and a few similar questions. The colleges and universities were asked to give various information about the courses, including "the method of instruction (large section, small section, etc.)." None of the data elicited by questions of this type is likely to give us insight into the existence, nature, or effects of imaginative and innovative methods of instruction, unless the "etc." in the last question noted proves to be a highly productive stimulant.

In much of the recent and current criticism of American education, both at the college and pre-college levels, the charge of "irrelevance" is commonly heard. A content analysis of the meaning of irrelevance in this context would probably reveal "dullness" as its chief component. There is little doubt that economics courses are notable for their dullness on many campuses. Professor Saunders expressed surprise that his survey showed "difficulty of subject matter" for students who took only the introductory economics course to be only 2.99 on a one-to-five-point scale where three is "average." I expect that he would not express surprise that "interest of subject matter" showed a mean score of 2.72, below average. If, as I am convinced, undergraduate college teaching is dull and unimaginative in general, then it is a serious indictment of any subject matter to be below the average of such teaching.

I believe that the three major requirements for a good teacher are that he be well informed on his subject, interested in his subject, and well informed on pedagogical methods. Most college teachers of economics are well informed about their subject; they are probably interested in their subject, although this may not show up in the introductory course; and they are dismally uninformed about pedagogical methods. Charisma saves a few professors from the dire consequences of their ignorance about pedagogy; with the rest, the students suffer through dull lectures, dull texts, and dull examinations.

The only hope for viability of undergraduate education in the future lies in drastic

changes from the currently typical methods. The changes should be, and hopefully will be, toward much greater variety in teaching-learning configurations, greater student autonomy, greater individualization of learning, use of a wider range of technological and social resources including community resources, and more concern with social applications of economic knowledge. I had intended to give examples of these hoped-for changes, but I found that Professor Lloyd, toward the end of this paper, listed a great many of the specifics that I would have mentioned.

I was, incidentally, greatly relieved by the comments that Professor Lloyd made toward the end of his paper. Near the beginning of his paper I thought that Professor Lloyd was pleading for education of blacks equal to education of whites. I was going to ask him whether he really felt this was sufficient, but he made clear that it is not—that he is calling for quality education for blacks well above that now given to white undergraduates. I quite agree. Designing good education for blacks may be more feasible at this point in our history than uplifting education for the general mass of students, most of whom are white. Hopefully, the exemplary programs for blacks will then give us leverage for raising the whole level of undergraduate education.

In summary, I congratulate Professor Saunders on a heroic effort. Within the confines of the orthodox content of introductory economics courses (as colorfully described by Saunders' Leftist and Rightist critics) and of my criticisms concerning the orthodox methods of teaching, the results are both promising and interesting. I recommend to Saunders a diligent search of the data for insights as to the effects of different teaching methods—something more imaginative and promising than large section versus small section—although I realize that such data may not exist. I also recommend, therefore, that the "next" study give particular attention to describing and assessing the effects of a variety of teaching-learning configurations—the more varied and radical, the better.

MARCIA L. HALVORSEN: My comment on Professor Lloyd's paper is quite brief. While I

agree with his major points and conclusions, there are some points that I believe need elaboration and greater emphasis when we discuss "Teaching Economics to Black Students." These points relate to the quality of instruction as against the quality of the students.

I agree that it is both necessary and important to recognize student deficiencies in such skills as reading, arithmetic and writing. But the all-too-often result—a condescending attitude on the part of the instructor—can, as Professor Lloyd points out, be disastrous. It can lead to a lower set of standards on the part of the instructor, a loss of respect on *both* sides of the podium and thus to a situation ripe for ineffective teaching and learning. The damage done cannot be overcome by audio-visual aids or the use of programmed materials.

Various studies in economic education over the last decade have pointed out that one element in the teaching process that seems to make a significant difference in the performance of students of economics is the ability of the individual instructor. I suggest that at this particular point in history it is probably the crucial element in the teaching of economics to black students.

To focus on students' skills (or lack of them) is to emphasize *their* ability to communicate with the instructor. However, teaching is a two-way process. The instructor must be able to communicate with them, as well. Only then is the instructor able to view their minds as "fires to be ignited" rather than as empty "receptacles to be filled." It is essential to effective teaching to recognize what the student comes to the classroom *with* as against what he comes *without*.

Why is the individual student in the class? What sacrifices have he and his family made for him to be there? What does he hope to get out of it? What are his views of the world and of the opportunities afforded him in it? What is his range of knowledge about, and his experience with, economic institutions? The instructor, particularly the white instructor, cannot presume to know the answers to these questions. They are neither obvious nor simple.

We can all agree that the subject matter of economics must be the same whether taught to black or to nonblack students. But I will ar-

gue that the differences among what black students come *with*, in terms of motivations for learning, perspectives and experiences, must be recognized and dealt with in the *effective* teaching of that same subject matter to blacks as against nonblacks. The burden is clearly on the instructor to identify and then deal with these differences.

Some of the differences act to the advantage of the black students. There is a clear realization on the part of black people, generally, that a good education is essential if they are to "make it" in this society. And today, it is apparent that a college degree is necessary to open many doors for black people. In fact, it really doesn't matter much *unless* one gets that degree.

Students at the black colleges and universities are much more free to benefit from good instruction insofar as they don't have the "hang-ups" associated with being a "guest" on the campus or a "guinea pig" or a "special case," as they are so often regarded on predominantly white campuses.

Let me speak for a moment about the black colleges, which, after all, account for the majority of black college students in this country today. Although there are few nonblack students on most of these campuses, the faculties are perhaps half black and half nonblack, and the proportion of nonblacks teaching economics is even higher. The white instructor is initially viewed with suspicion, as the students try to figure out why he or she is there. Their suspicions are justified. Too many white instructors come with a missionary's vision of doing good works. These people, of course, cannot successfully teach black students. The students are there for an education, not a conversion.

Other white instructors, who think of themselves as liberals, also fail to teach black students successfully. They live in all-white suburbs, expend little effort in learning about the black experience, and, in general, remain aloof from the concerns of the students. They keep busy filling empty receptacles; the student dutifully hand back the contents of the receptacles on examinations; and everyone goes away relatively unscathed by the whole experience.

It takes much more than that. It takes a strong dedication to teaching. It takes intensive work with individual students on the part of the instructor. Hours must be spent listening to them, tutoring them in the basic skills, grading and then discussing with them their drills and exercises and problem sets. I agree with Charles P. Kindleberger, who worked with us for a year recently in the Atlanta University Center. He said this in a talk about a year ago:

A naively optimistic view runs to the effect that teaching machines at the lower schools and colleges can relieve our society from the labor-intensive task of teaching students by substituting labor-saving capital in the form of videotapes, films, programmed software and hardware, and the like, for the teacher, or at least reducing the teacher-student ratio. I am deeply sceptical. Informed conclusions wait on more evidence, but my intuition tells me that there is no substitute for the motivation provided by the devoted teacher, interested in the student as a person and in his welfare, ready to put out long hours in talking to students, correcting his tests and homework—and not on a true-false, yes-and-no basis—and dealing with him face to face. This may be provided from a Big Brother one day by mechanical means. I do not see it as currently available except through human beings.

It is important to remember that economics is a nontraditional field for black people. The instructor of economics must relate as a model to the student. He must have high academic standards for himself as well as for his students. He must show them how the substance of economics is both relevant and important to the understanding of the students' past, present and future in this country.

The discipline of economics has only recently been pursued by black students. Only in the last decade have many black colleges and universities begun to expand and upgrade their economics curriculum. Students are suspicious of the only recent overtures being made to them to enter the field. Again, the burden is on the instructor to help cut through the wedge of hypocrisy that exists in the way

colleges and graduate schools recruit black students and in the way businesses recruit black employees. He must point out the real and exciting opportunities that, indeed, do exist and are expanding for black students.

To perform all of these functions is not an easy job, whether you are a black or a white instructor, and whether you teach in a predominantly black college or a predominantly white college. I suspect that predominantly white institutions should take to heart Professor Sowell's recent article in the *New York Times Magazine*, dealing with how colleges are skipping over competent blacks to admit "authentic" ghetto types. It is becoming increasingly clear that these institutions are recruiting black students for the wrong reasons and that they are not expending the kinds of effort necessary to acquire the ability to successfully teach ghetto students. And they are having all kinds of trouble. Perhaps they should concentrate on recruiting middle-class, high-quality students (as reflected on S.A.T. scores). They are better equipped to teach these students successfully.

In conclusion, with respect to teaching economics to black students, I have focussed my brief remarks on some of the issues connected with the quality of instruction and/or the qualities of a good instructor. Our training in the profession prepares few of us to teach. We "pick it up" somewhere along the way. The current excursions being made in institutions around the country in the teaching of economics to black students put the spotlight on many of the deficiencies in our own capabilities as teachers. The adjustments needed to bring about successful and effective teaching lie heavily on those of us who profess to be teachers of economics.

CAMPBELL R. MCCONNELL: I have been asked to react to Professor Castro's paper and am pleased to report that there is much to agree with and commend. His argument that many academics have used the medium of television badly is unquestionably accurate. His emphasis upon a mixture of television and live discussion is undoubtedly prudent. Furthermore, given present knowledge about televised instruction—that the potential cost

economies are great, that students can learn from TV as well as they can from more conventional methods, and that students find televised instruction uninteresting and regard it as an inferior substitute for live instruction—one is inclined to be enthusiastic about a dialogue approach to the use of television.

Other things being equal, it is difficult to resist such educational techniques which put a high priority on student interest and involvement. But the catch is: does *ceteris paribus* hold? Or are we perhaps faced with a pedagogical Phillips Curve—let us call it a Castro Curve—which poses a trade-off between interest and "intellectual interaction," on the one hand, and the conveying of an accepted body of principles, on the other. Professor Castro seems to be saying that dialogue tapes provide us with a better Castro Curve, a better menu of choices whereby the student's grasp of economic analysis and his interest in the subject matter may be simultaneously enhanced in comparison to alternative methods. His crucial assumption is that dialogue tapes will fan latent student interest in socioeconomic problems and "disciplinary sophistication will follow because our students will want it. . . ."

Whether this assumption is valid may turn upon the issue of precisely how and the extent to which dialogue tapes are used. Professor Castro leaves us in the dark on this point, so we must speculate. If dialogue tapes are used on an occasional basis and followed by the application of economic principles in assessing the major points of controversy and a delineation of the reasons why economists disagree, then a better Castro Curve may result. But, if dialogue tapes are used with great frequency as the core instructional mechanism for the introductory course and followed by an unstructured, student-centered discussion, the results might be less desirable.

My limited observation of dialogue or controversy approaches suggests that the ultimate goal of each participant is to prevail—to gain general acceptance of one's position, to win the debate. I note with dismay that the customary constraint of reasoning in accordance with the canon's of elementary logic is frequently abandoned in such encounters and that one is often inclined to neglect, or misap-

ply, the principles of one's own discipline should they be awkwardly at odds with one's views. Advocacy dominates analysis; ideology prevails over intellectual integrity. In the heat of the debate the spoils frequently go to the articulate, rather than to the accurate. This, I think, is a potentially serious limitation of the dialogue approach. Instead of setting the stage for the uncovering of basic principles, the dialogue approach may serve to reinforce preconceptions and biases, making the task of digging out the relevant theory seem redundant and irrelevant to students. Given the disconcerting body of evidence that many traditionally-conceived principles courses are now quite ineffective in conveying an understanding of even the most basic analytical tools, I would view this possible outcome with considerable alarm. Stated differently, one can plausibly argue that some degree of disciplinary sophistication—some grasp of elementary analytical tools—must *precede* any meaningful examination of controversial issues.

A related point is that a dialogue approach may pose the danger of undermining student belief in the integrity of the discipline. The emphasis of the dialogue approach is upon disagreement and discord, wherein actually the degree of agreement in economics is remarkably high. While Professor Castro may be correct in lamenting that testing and programmed instruction create a "single 'right' answer bias," it would be equally unfortunate if dialogue tapes were to leave students with the impression that there are no answers at all.

Let me close on a more positive note. I would speculate that much of the ineffectiveness—real and imagined—of the principles course stems from its traditional rigidities. An economics department typically decides unilaterally how it wants to offer principles, and

all students are funnelled through the same course. My suggestion is that we might profitably view the problem of effectively and efficiently teaching the principles course as an allocational problem. We have made progress in recent years on the supply side; that is, we have begun to experiment widely with various pedagogical alternatives and to measure their effectiveness. Televised lectures, programmed courses, lectureless courses, gaming and computerized instruction, and now dialogue tapes all come to mind. At many institutions the student-customer has two or more pedagogical options from which to choose. In my judgment this is a most desirable development. But perhaps we now need to pay greater attention to the demand side. Students need more information about their own learning capacities under various instructional systems if they are to exercise rational choice. The next major step in economic education may entail efforts to identify which students will learn most effectively and efficiently with a given alternative. Which students learn best by passively viewing televised lectures? By working through programmed materials? By actively participating in problem-oriented seminars? By reacting to dialogue tapes? My point is that the student-customer not only needs to be confronted with a variety of instructional options, but he also needs some information with respect to which option is most appropriate for him. If a Castro Curve exists—if there is a trade-off between stimulating student interest in economics and transmitting a firm grasp of the analytical tools of economics—that curve (like the Phillips Curve) may have microeconomic roots. Perchance the allocational approach which I have sketched may lead to a better curve than is now available.

EDUCATIONAL PRODUCTION RELATIONSHIPS

Education, Technology, and the Characteristics of Worker Productivity*

By HERBERT GINTIS

Harvard University

Economists have long noted the relationship between the level of schooling in workers and their earnings. The relationship has been formalized in numerous recent studies of the rate of return to schooling and the contribution of education to worker productivity. Almost no attempt has been made, however, to determine the mechanism by which education affects earnings or productivity. In the absence of any direct evidence, it is commonly assumed that the main effect of schooling is to raise the level of cognitive development of students and that it is this increase which explains the relationship between schooling and earnings. This view of the schooling-earnings linkage has provided the conceptual framework for studies which seek to "control" for the quality of schooling through the use of variables such as scores on achievement tests and IQ. [26, 46]

* This research was supported by a grant from the Carnegie Foundation of New York to the Center for Educational Policy Research, Harvard Graduate School of Education, and a separate grant from the Social and Rehabilitation Service, U. S. Dept. of Health, Education, and Welfare. Special thanks are due to Stephan Michelson, Christopher Jencks and Samuel Bowles, who aided the research through its various stages, and the Union for Radical Political Economics collective at Harvard, whose members helped broaden its scope. This work is part of a larger project in process jointly with Samuel Bowles, on the political economy of education. The arguments presented herein have been significantly abridged, in conformance with stringent space limitations. Requests for amplification of the material should be addressed to the author.

The objective of this paper is to demonstrate that this interpretation is fundamentally incorrect. It will be seen that rejection of the putative central role of cognitive development in the schooling-earnings relationship requires a reformulation of much of the extant economic research on education, as well as a radical rethinking of the normative bases of the economics of education in particular, and neo-classical welfare economics in general. In Section I, I will present data to suggest that the contribution of schooling to worker earnings or occupational status cannot be explained by the relationship between schooling and the level of cognitive achievement. Indeed, the data there introduced strongly suggest the importance of noncognitive personality characteristics which have direct bearing on worker earnings and productivity. In Section II, I will give substantive content to the relevant personality variables operative in the relationship between education and earnings. With the theoretical literature on the personality requisites of adequate role-performance in a bureaucratic and hierarchical work-environment as a frame of reference, I will sketch some mechanisms through which schools affect earnings. This involves scrutinizing the social relations of education and the pattern of rewards and penalties revealed in grading practices. I will argue that the authority, motivational, and interpersonal relations codified

in the "social structure" of schools are closely similar to those of the factory and office. Thus a path of individual personality development conducive to performance in the student's future work roles is facilitated. Further, I will show that the grading structure in the classroom reflects far more than student's cognitive attainment, by affording independent reward to the development of traits necessary to adequate job performance.

If my basic thesis is correct, much of the existing body of economic literature on schooling must be reconsidered. First, we must redefine the concept of "quality" in education, particularly in studies of the determinants of earnings which have thus far relied on measures of cognitive development as the sole measure of educational quality (e.g., [9]). Second, the extensive body of research on "educational production functions" is seen to lack economic relevance, since the dependent variables in most of these studies have been restricted to measures of cognitive achievement [4, 9, 28]. Third, the extensive body of literature on resource allocation in schooling—extending from planning model to rate of return studies—requires reformulation. The normative base of these studies requires that the mechanism by which schooling contributes to earnings operates independently from the character structures of the individual students. That is, they assume that the process of schooling does not affect the tastes and personalities of the future workers being processed for higher productivity. Yet the data below strongly suggest that the economic productivity of schooling is due primarily to the inculcation of personality characteristics which may be generally agreed to be inhibiting of personal development. The "economic productivity" of schooling must be measured against an "opportunity cost" reflected in the development of an alienated and repressed labor force. Fourth, the

above point is simply a special case of a more general problem in neoclassical welfare economics. Our analysis shows that taste and personalities are not determined outside the economic system, but are rather developed as part of the economic activities about which social policy is to be made. Thus the main theorems of welfare economics, being based on the independence of individual preferences from the structure of economic institutions [19], fail.

While our evidence suggests the reformulation of much of the existing work in the economics of education and welfare theory, it may also provide resolutions to some of the outstanding anomalies that have arisen in recent years. For example, a number of studies have shown very low monetary returns to the education of lower class people and blacks in the U. S., even with the level of cognitive development taken into account [51]. These results are readily explained by our model, where they likely result from the failure of schooling to inculcate the required noncognitive personality traits in the observed groups. Moreover, it is often found, both in the U. S. and in underdeveloped countries, that the economic return to vocational schooling is quite low [45]. This is especially surprising in that vocational education dwells exclusively on the supposedly "economically relevant" content of schooling. Our interpretation renders the finding of low economic returns of vocational schooling understandable, in view of its misplaced emphasis on the "skill content" of schooling, and a corresponding underemphasis on the broader socialization function involving the generation of a disciplined, obedient, and well-motivated workforce. Lastly, recent years have seen the revival of so-called "genetic" theories to explain the pattern of racial and social class inequalities [15, 32]. Neither proponents nor opponents of this view seem seriously

to have questioned the importance of cognitive ability in occupational status and earnings, but have restricted their considerations to the narrow question of "heritability" of intelligence, in the naive view that IQ lies at the heart of economic success. Our results would indicate that this debate is close to irrelevant, save at the very extremes of the "ability" distribution.

*The Cognitive Element in Schooling's
Contribution to Worker Earnings*

The "market value" of a worker depends on a certain array of personal characteristics—cognitive, affective, and ascriptive.¹ The bulk of modern sociological theory affirms the minor importance of ascriptive traits in the general allocation of social roles and status positions, at least among white, male Americans. Thus we may take the individual traits generated or selected through schooling, insofar as they relate to the augmentation of worker earnings, as predominantly cognitive and affective. Hence we propose to test the adequacy of two polar "ideal type" models—the Cognitive and the Affective. According to the Cognitive Model of education's contribution to worker earnings and occupational status (Y), the variable E (years of education, corrected for differences in "quality" in the form of physical, teacher, peer-group, and content resources) is a proxy for a set of cognitive achievement variables A (e.g., reading speed, comprehension, reasoning ability, mathematical or scientific achievement). According to the Affective Model, on the

other hand, E is a proxy for a set of relevant personality variables P . Using a linear regression model to capture the income- and occupational-status-generating process, a test of the Cognitive Model is particularly immediate. If Y is a measure of income and/or occupational status, then the "contribution of education" can be interpreted as the beta coefficient in the regression

$$(1) \quad Y = a + b_{YE}E + u.$$

If the Cognitive Model is correct, then in the extended regression

$$(2) \quad Y = a + b_{YA \cdot E}A + b_{YE \cdot A}E + u,$$

we expect $b_{YE \cdot A} = 0$. That is, introducing achievement variables into the restricted regression (1) reduces the contribution of E by 100%. If, on the other hand, the Affective Model is correct, and if A and P are related only through their common dependence on E , then $b_{YE} = b_{YE \cdot A}$, and the reduction in the contribution of E is 0%.

Clearly we have divergent implications, empirically testable by available data. Appendix I exhibits the results of many studies, including measures of Y , A , and E , comprising all investigations the author has come upon.² These studies, despite their divergent measures of relevant variables and use of distinct sample populations, show two broad uniformities: (a) The reduction in the coefficient of E due to the introduction of achievement variables is much closer to zero than to 100%—the actual range is 4% to 35%; and (b) the increase in explained variance is negligible—i.e., less than 5% of explained variance. At first glance, these studies provide strong support for the Affective Model, and indi-

¹ By 'cognitive characteristics' we mean individual capacities to logically combine, analyze, interpret, and apply informational symbols. By 'affective characteristics' we mean propensities, codified in the individual's personality structure, to respond in stable emotional and motivational patterns, to demands made upon him in concrete social situations; and by 'ascriptive characteristics', we mean such non-operational attributes as the individual's race, sex, caste, religion, social class, eye color, geographical region, etc.

² These are restricted to U. S. samples, predominantly white, male, average mean intelligence. Also, I have indiscriminantly mixed "achievement" and "intelligence" measures of A . The results of these investigations are strikingly similar for both measures, so their synthesis presents no problems for the purposes of our investigation.

cate that cognitive development is not the central means by which education enhances worker success.

Two possible objections to this analysis, however, induce us to expand the model. On the one hand, there may be relations between P and A beyond their common dependence on E . This might occur either through a direct relation between P and A (i.e., income-relevant personality traits facilitate the acquisition of cognitive achievement—see below) or because both A and P depend on variables not included in our simple equations, such as genes and social class. In either case, the reduction in b_{YE} through the inclusion of A would exceed that of a properly specified model, and the outcome would be even more in favor of the Affective Model, in that this model becomes compatible with the observed small but significant reduction (see below). On the other hand, one might hold that while theoretically the introduction of A into the income-education regression should reduce the coefficient of E to zero, in fact, both A and E are so subject to errors in measurement that the results are significantly altered in practice.³ In particular, if E is really a “proxy” for A , but A is measured with significantly greater error than E , the latter becomes a more reliable indicator of achievement than A itself.

We shall take these objections seriously and introduce a more extensive model, including important background measures of abilities (I) and social class (S). Moreover, we shall allow for the “observed” measures of E and A , which we denote by

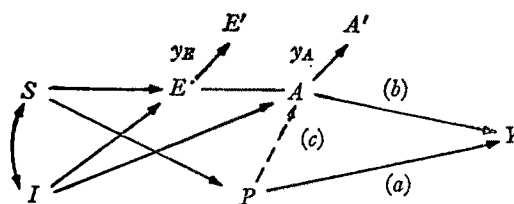


FIGURE 1

E' and A' , to include an element of random error, so that $r_{EE'} = y_E$ and $r_{AA'} = y_A$. It will be assumed that all errors are uncorrelated, so that recursive regression analysis may be applied.⁴ The recursive schema is shown in Figure 1, where the elimination of path (a) corresponds to the Cognitive Model, and the elimination of path (b) to the Affective Model (the dotted arrow (c) will be discussed later).⁵ Here P is treated as a “hypothetical variable” [29], in that we shall not specify its content in this Section. As part of the Affective Model, however, we shall assume that education (E) and social class (S) are *important* elements in the determination of P , and that education has at least as great a direct importance as social class.

We must now recalculate the “expected” fall in b_{YE} due to the introduction of cognitive variables, based on this larger model, and with y_A and y_E as parameters. It will be shown that for all reasonable values of these parameters—and for many unreasonable as well—the Affective Model predicts far more accurately than the Cognitive. The Affective Model in its crude form tends to *underpredict* (predicted reduction = 0%, while actual re-

⁴ The statistical techniques of recursive regression, or “path analysis”, are described in [14].

³ By “errors of variables” we include more than simple test reliabilities and validities in reportage, but the larger errors arising from an incomplete or a partially misdirected measuring instrument. Thus to measure E by “years of educational attainment” introduces errors because this measure abstracts from the *quality* of schooling. Similarly, a measure of “achievement” may inherently capture only a portion of the “theoretical” variable.

⁵ It might be asked why certain paths have been *a priori* excluded from Figure 1. That Y depends directly only on A and P follows from our exclusion of the influence of ascriptive traits and further from the very definition of A and P themselves; insofar as Y depends on I , I should be included among the variables A , and insofar as Y depends on traits involved in S , S must be included directly in P . Studies show moreover, that the direct path from S to A is negligible [29, 40].

duction=4%–35%). Hence in reestimating its predictions, we shall remain conservative by always underestimating the Affective Model's predicted reduction in b_{YE} when A is introduced. Manipulation of the normal equations in two-variable regression ([33], p. 61) gives

$$\frac{b_{YE \cdot A}}{b_{YE}} = 1 - \frac{r_{EA}}{1 - r_{EA}^2} (z - r_{EA})$$

where $z = r_{AP}/r_{EP}$. In terms of the imperfectly measured variables E' and A' , this clearly becomes

$$(4) \quad \frac{b_{YE' \cdot A'}}{b_{YE'}} = 1 - y_A^2 \frac{r_{EA}}{1 - y_A^2 y_E^2 r_{EA}^2} \cdot [z - r_{EA}] + [1 - y_E^2] r_{EA}.$$

To estimate this equation, taking y_E and y_A as parameters, we require estimates of r_{EA} and $[z - r_{EA}]$. A conservative assessment of the Affective Model requires that we choose a small value for r_{EA} , since the larger is r_{EA} (holding $z - r_{EA}$ constant), the larger the predicted reduction in $b_{YE'}$. Empirical measures [11, 13, 26] show $r_{E'A'} \approx .6$ to $.7$. Since $r_{EA} = r_{E'A'}/y_A y_E$, the assumption of significant errors in variables pushes r_{EA} quite high (in terms of the assumptions of Table 1, even above unity).

TABLE 1

y_A^2	y_E^2	Reduction in $b_{YE'}$, percentage		
		Model A $p_{AP}=0$	Model A $p_{AP}=.12$	Cognitive Model
1.00	1.00	00%	24%	100%
0.75	0.85	08%	19%	73%
0.60	0.70	11%	18%	49%
0.50	0.70	09%	15%	40%
0.80	0.80	12%	22%	72%
0.80	0.70	16%	27%	71%
0.70	0.60	17%	26%	62%

However, there is reason to believe the error in measuring A is *not* independent of E (e.g., through the conceptual variable

"test-taking ability," which might increase with level of education) so our underestimate of r_{EA} will take the form of *not* correcting for measurement error; that is, we assume E accounts for about 50% of the variance in achievement.

Similarly, we shall settle for an underestimate of $[z - r_{EA}]$. Abstracting from p_{AP} , and using the fundamental theorem of path analysis [14], we have

$$(5) \quad r_{AP} - r_{AE} r_{PE} = p_{AI} p_{PS} \{ r_{EA} [1 - p_{EI}^2 - p_{EA}^2 - r_{EA} p_{EI} p_{ES}] - p_{EI} p_{ES} \}.$$

Thus in general we have

$$(6) \quad [z - r_{EA}] = r_{PE}^{-1} \{ p_{AP} + p_{AI} p_{PS} \cdot [r_{EA} (1 - R_E^2) p_{EI} p_{ES} (1 - r_{EA}^2)] \}.$$

Using figures from [11, 40], we find the highest estimate of the second term on the left-hand side of (6) to be $p_{PS}(.05)/r_{PE}$. Since

$$(7) \quad r_{PE} = p_{PE} + r_{ES} p_{PS},$$

we have

$$(8) \quad 1 = (p_{PE}/r_{PE}) + r_{ES} (p_{PS}/r_{PE}).$$

Since $r_{ES} \geq 0.6$ [15], we have $(p_{PS}/r_{PE}) < 1.7$. But assuming the direct link between E and P to be strong (an assumption of the Affective Model), this estimate is seen to be significantly inflated. Indeed, if S and E are roughly equal in their direct linkage with P , (p_{PS}/r_{PE}) is significantly less than unity. At any rate, the elimination of the second term in (6) will not bias our results greatly against the Affective Model.

In treating r_{PE} , we shall again settle for an underestimate of $[z - r_{EA}]$. We have

$$(9) \quad P = p_{PE} \cdot E + p_{PS} \cdot S + p_u U_p$$

where U_p is the contribution to P outside the model. If we then write $p_{PS} = p_{PE} I_a$, we

find that

$$(10) \quad r_{EP} = R_P / \sqrt{[(1 - r_{ES}^2)/(a + r_{ES}^2)] + 1},$$

where R_P is the "proportion" of the variance of P explained within the model. Moreover, taking $r_{ES} = .6$ [1, 13], the denominator on the right side of (10) declines from 1.24 to 1.00 as " a " passes from $\frac{1}{2}$ to infinity. Thus we can safely take

$$(11) \quad R_P^{-1} < r_{EP}^{-1} < (1.2)R_P^{-1}.$$

Returning to (6) and (4), we find that the Affective Model implies a reduction in $b_{YE'}$ with the following upper bound:

$$(12) \quad \frac{b_{YE' \cdot A'}}{b_{YE'}} < 1 - y_A^2 \frac{r_{EA}}{1 - y_A^2 y_{E'}^2 r_{EA}^2} \cdot \{p_{AP} R_P^{-1} + [1 - y_E^2] r_{EA}\}.$$

Moreover, this upper bound is probably a good approximation as well, so (12) can be treated as the "prediction" of the reduction in $b_{YE'}$ by the Affective Model.

We shall test equation (12) using two estimates of p_{AP} . First, we shall assume there is no direct relation between A and P , so $p_{AP} = 0$. Second, we shall assume a small direct relation, taking (arbitrarily) $p_{AP} = .12$.⁶

The corresponding analysis for the Cognitive Model's prediction of the reduction in $b_{YE'}$ follows from a similar but simpler derivation. We find

$$(13) \quad \frac{b_{YE' \cdot A'}}{b_{YE'}} = 1 - \frac{y_A^2}{1 - y_A^2 y_{E'}^2 r_{EA}^2} \cdot [1 - y_E^2 r_{EA}^2].$$

Table 1 illustrates our predictions for alternative hypotheses as to the validities of A and E , and with our alternative

⁶ Theoretical evidence of such a direct relationship can be found in [44]. In addition, I shall suggest below that a central trait developed through schooling and relevant to job adequacy is "motivation according to external reward," which is clearly conducive to higher levels of cognitive achievement.

assumptions concerning p_{AP} .⁷ If the empirically derived reductions shown in Appendix I are correct, the Cognitive Model must be fairly decisively rejected. Moreover, the Affective Model "predicts" with the proper order of magnitude. Of course, the latter's validity can only be ascertained when a correct specification of the variables P are obtained. A preliminary attempt in this direction will be presented in the following Section. Roughly, if education does contribute to earnings, and if this contribution cannot be accounted for in terms of cognitive variables, it is reasonable to expect the noncognitive traits rewarded through grading and promotion, hence presumably integrated into the student's personality, to do the job.⁸

The Structure of Social Relations and the Pattern of Rewards in Schooling

The cogency of the Affective Model depends in the last instance on the quantitative specification of the personality variables P . Ideally, this would involve isolating a fixed set of measurable traits, exhibiting their concordance with level of occupational status or income, showing their correlation with years of school, and describing the mechanisms by which schools generate them. In this section, our

⁷ As previously noted, Table 1 is based on the assumption that $r_{EA}^2 = .5$. The reader can verify that these results are quite insensitive to alternative specifications of r_{EA}^2 .

⁸ Approaches to the sources of worker productivity other than that followed in this Section are available. For instance, noting that the demand and supply of educated labor have increased in step in recent years, in that relative wages have not significantly changed [2, 22], we may ask if the rise in demand can be accounted for in terms of cognitive variables. The cognitive requirements for jobs included in [47] exhibit high reliabilities [16], and even tend to overestimate these requirements [38]. An analysis of this body of data [48, 39] shows cognitive demands requiring a total increase of 0.44 years of education per worker, between 1940 and 1960, whereas the actual increase in level of education is several times this value. Thus the demand for "educated labor" must include significant noncognitive components.

aim will be considerably less ambitious. We shall outline a set of traits held by long sociological tradition [37, 49, 50] as requisite for adequate job functioning in production characterized by bureaucratic order and hierarchical control. We then show that schooling is conducive to the development of these traits in students.⁹

We shall focus on two aspects of schooling central to the patterning of personality development. First, the *structure of social relations* in education—including sources of motivation, authority and control, and types of sanctioned interpersonal relations—by requiring the student to function routinely and over long periods of time in role situations comprising specific expectations on the part of the teacher, other students, and administrators, tends to elicit uniformities of response codified in individual personality [5, 12]. Second, the system of grading, by rewarding certain classroom behavior patterns and penalizing others, tends to reinforce certain modes of individual response to social situations. According to any of the variants of behaviorist psychology, this *pattern of reward* will educe the corresponding pattern of personality traits in the students. Part of the myth of liberal education is that, however important the teacher's expectation may be in eliciting student performance, his actual *assessment* and *grading* of this performance depends only on concrete, observed cognitive attainments. Yet

⁹ In this paper we shall treat only those required traits which are *common* to all levels in the hierarchy of production, and are inculcated in most schools on all levels. Actually the personality requisites of job adequacy no doubt vary from level to level within the hierarchy of production, and different levels of schooling (e.g., grade school, high school, junior college, college) likely reflect these differential needs. Moreover, within a particular educational level, we would expect different types of schooling to subsist side by side (e.g., ghetto, working-class, and middle-class-suburban high schools), reflecting the differential positions in the production hierarchy that its students are destined to fill. These complications, however important, cannot be treated here.

studies show that cognitive variables never account for more than 30% of the variance in grade point average (37). In addition, many studies illustrate the importance of specific personality measure in prediction grades [17, 25, 30], although these by and large correct inadequately for actual achievement levels.

Before attempting a systematic assessment of the body of empirical information on the effect of social structure and pattern of reward in schooling on personality development, I should like to present two studies [19] illustrating the breadth and counter-intuitive nature of the process of grading, in which explicit measures of cognitive achievement are available. First, an analysis based on data collected on 649 upper-ability senior-high school males [31] (National Merit Scholarship Finalists) shows no value of any combination of five achievement variables (College Entrance Examination Board: SAT-Math, SAT-Verbal, Scientific Performance, Humanities Comprehension, Scientific Comprehension), despite significant variance in these achievement measures and in grades within the sample. Of some 65 additional personality variables, two—"Citizenship-Teacher's Rating" (CitT) and "Drive to Achieve-Student Self Rating" (DrA)—have greatest power to predict GPA, with $p < .001$. This example illustrates (a) since these traits are not rewarded *through* their contribution to achievement, that teachers grade *independently* on the basis of personality; (b) since DrA is rewarded, that subjective motivation is taken into consideration in grading; (c) since CitT is positively rewarded and can be interpreted as "conforming to the dominant role-structure" of the school, that grading reinforces the student's personality development through participation in the particular structure of social relations in schools, and; (d) while grades depend on achievement in general, when 'ability' is con-

trolled, little additional effect of achievement can be detected, so the *subjective experience* of an individual student (who of course cannot control his intelligence) is that grades depend primarily on affective

behavior. In this study, the pattern of reward is no less reflected in the remaining personality traits (Tables 2-4). Thus Table 2 shows that students are uniformly penalized for creativity, autonomy, initiative,

TABLE 2—PERSONALITY VARIABLES CORRELATED WITH GPA CORRECTED FOR ACHIEVEMENT, CIT, AND DRA (SIGNIFICANCE LEVELS IN PARENTHESIS)

<i>Positively Rewarded:</i>	SAT-Math (15%)
Perseverance (1%)	Scientific Comprehension (15%)
Good Student (1%)	<i>Negatively Rewarded (Penalized):</i>
Self-Evaluation (5%)	Independence-Self-Reliance (1%)
Popular (5%)	Initiative (5%)
Acceleration of Development (5%)	Complexity of Thought (5%)
Mastery (5%)	Originality (Barron) (6%)
Control (6%)	Originality (11%)
Status (11%)	Independence of Judgment (13%)
Popularity (TR) (13%)	Creative Activities (13%)
Suppression of Aggression (14%)	Curious (14%)

TABLE 3—CORRELATIONS OF VARIOUS PERSONALITY TRAITS WITH CIT

<i>Positively Rewarded:</i>	Mastery (15%)
Deferred Gratification (1%)	Initiative (15%)
Perseverance (1%)	<i>Negatively Rewarded (Penalized):</i>
Control (1%)	Cognitive Flexibility (1%)
Popularity (1%)	Complexity of Thought (1%)
Social Leadership (1%)	Originality (Barron) (1%)
Good Student (Parent value) (1%)	Sense of Destiny (1%)
Self-Evaluation (1%)	DRS-Creativity (1%)
Scientific Comprehension (1%)	Independence of Judgment (5%)
Intellectuality (1%)	Independence-Self-Reliance (10%)
Esthetic Sensitivity (5%)	Curious (15%)
Suppression of Aggression (5%)	Self-Confidence (15%)
Comradeship-Sharing (5%)	Verbal Activity (15%)
SAT-Math (10%)	
Artistic Performance (10%)	

TABLE 4—CORRELATIONS OF VARIOUS PERSONALITY TRAITS WITH DRA

<i>Positively Rewarded:</i>	Initiative (1%)
Self-Evaluation (1%)	Status (1%)
Perseverance (1%)	Breadth of Interest (5%)
Deferred Gratification (1%)	SAT-Math (5%)
Originality (1%)	Scientific Performance (5%)
Independence (1%)	Verbal Activity (5%)
Responsibility (1%)	Conformity (10%)
Control (1%)	<i>Negatively Rewarded (Penalized):</i>
Artistic Performance (1%)	Cognitive Flexibility (1%)
Creative Activities (1%)	Complexity of Thought (1%)
Sense of Destiny (1%)	Originality (1%)
Popularity (1%)	Independence of Judgment (1%)
Social Leadership (1%)	SAT-Verbal (5%)
Good Student (1%)	
Mastery (1%)	

tolerance for ambiguity, and independence, even after correcting for achievement, CitT, and DrA, and rewarded for perseverance, good student values, and other traits indicative of docility, industry, and ego-control. Moreover, the content of CitT is exhibited in Table 3, showing a similar pattern of evaluative behavior on the part of the teacher, especially in the penalized traits (no doubt as a result of the objective needs of 'classroom control' in the typically-structured school, rather than his personal value-preferences). Lastly, Table 4 shows that DrA is associated with the same pattern of penalized traits, while the rewarded traits exhibit two separate dimensions: on the one hand, high DrA may involve conformity with classroom norms, and on the other, to their rejection in favor of autonomous personal development—hence the appearance of Artistic Performance, Creative Activities, Self-Confidence, Initiative, Self-Assurance, Breadth of Interest as associated with DrA.¹⁰

The National Merit Scholarship study is weak in two respects. First, it deals with only one ability grouping, and second, it aggregates over diverse study-areas—natural science, social science, humanities, language, etc. A similar result can be derived, however, from a path analysis the author has fit to data supplied by Cline [7], covering 114 high school seniors of varying ability, in the specific area of natural science performance. This data-source includes a measure of intelligence, three creativity measures, achievement level in science, a teacher rating of the student's "science potential," and average science grades over the three years of high school. Path analysis [19] indicates that

over the broader ability spectrum of students: (a) teacher attitudes are the major determinants of grades; (b) "achievement" is only one determinant of teacher attitudes, and hence of grades received; (c) intelligence is directly rewarded in terms of grades, beyond its contribution to actual achievement, whereas many other equally important determinants of achievement (e.g., "creativity") are in no way rewarded; (d) the direct effect of actual achievement on teacher attitudes is statistically insignificant.

The bulk of existing studies are compatible with these results, and hence tend to lend credence to the Affective Model. Moreover, these studies show that both structure and pattern of reward in schooling conform to the requisites of adequate job-performance in bureaucratically structured and hierarchically organized enterprise [37, 49, 50]. We can organize this discussion around four types of personality requisites—"Subordinacy," "Discipline," "Supremacy of Cognitive over Affective Modes of Response," and "Motivation according to External Reward."

Subordinacy. "The principle of hierarchical authority . . . is found in all bureaucratic structures . . . (in a) firmly ordered system of super- and sub-ordination." [49]. Subordinacy and proper orientation to authority are induced through the strict hierarchical lines—administration-teacher-student—of the school. As the worker relinquishes control over his activities on the job, so the student is first forced to accept, and later comes personally to terms with, his loss of autonomy and initiative to a teacher who acts as a superior authority, dispensing rewards and penalties. That proper subordinacy is a factor in grading as well is dramatized in our National Merit Scholarship study. It is supported by Gough [24], where "over-achievers" (students whose grades exceed

¹⁰ This divergence replicates [18]. Here, as throughout this Section, space limitation forbids adequate explanation of the content of these personality measures. Their precise content can be found in the corresponding sources, or in [19].

that predicted by their IQ) are marked by their teachers as "appreciative," "cooperative," and "reasonable," while "under-achievers" are deemed "dissatisfied," "pre-occupied," "rebellious," and "rigid." Striking additional support is found in [6] (see [19]).

Discipline. Weber emphasizes, "organizational discipline in the factory is founded upon a completely rational basis . . . the optimum profitability of the individual worker is calculated like that of any material means of production. On the basis of this calculation, the American system of 'Scientific Management' enjoys the greatest triumphs in the rational conditioning and training of work-performance . . . the psycho-physical apparatus of man is completely adjusted to the demands of the outer world . . ." The extension from production on simple factory lines to bureaucratic organization both conserves and expands this need. In Merton's words [37], "bureaucratic structure exerts a constant pressure . . . to be 'methodical, prudent, disciplined.' . . . The bureaucracy . . . must attain a high degree of reliability of behavior, and unusual degree of conformity with prescribed patterns of actions. Hence the fundamental importance of discipline. . . ." Discipline is reflected in the educational system where regularity, punctuality, and quiescence assume proportions almost absurd in relation to the ostensible goals of "learning." Thus Gough [24] finds overachievers consistently rewarded for being "dependable," "reliable," "honest," and "responsible," (teacher ratings), and Gebhart and Hoyt [17] find them rewarded for "Consistency" (Edwards Personal Preference Inventory); our National Merit Scholarship study shows deferred gratification, perseverance, and control as central elements in the teacher's assessment of "citizenship," a highly rewarded trait. Dramatic conforma-

tion of discipline as independently rewarded through grades is supplied by a brilliant series of studies by Smith [42, 43, 44]. Noting that personality inventories suffer from low validities due to their abstraction from real-life environments, and low reliabilities due to the use of a single evaluative instrument, Smith turned to student peer-ratings of 42 common personality traits, based on each student's observation of the actual classroom behavior of his classmates. Factor analysis allowed the extraction of five general traits, stable across different samples and national cultures. One of these, a discipline factor which Smith calls "Strength of Character"—including such traits as "not a quitter," "conscientious," "responsible," "insistently orderly," "not prone to day-dream," "determined-persevering,"—exhibits three times the contribution of R^2 in predicting post-high-school performance than any combination of thirteen cognitive variables.

Cognitive vs. affective modes of response. Occupational roles have been characterized as requiring an upgrading of cognitive demands. Yet the contribution of schooling to job-adequacy may be more accurately described as evincing a *cognitive mode*, and suppressing an affective mode, of reacting to bureaucratic situations. That bureaucratic order requires the dominance of cognitive modes of response has been emphasized by Weber [49]: "bureaucratization . . . very strongly furthers the development of 'rational matter-of-factness' . . . Its specific nature, which is welcomed by capitalism, develops the more perfectly the more the bureaucracy is 'dehumanized,' the more completely it succeeds in eliminating from official business love, hatred, and all purely personal, irrational, and emotional elements which escape calculation." More recently, Keniston [35] notes, "The preferred technique of

technology involves two related principles: that we give priority to cognition, and we subordinate feeling. . . . Thus feeling as a force of independent value—all of the passions, impulses, needs, drives, and idealisms which in some societies are the central rationales of existence—are increasingly minimized, suppressed, harnessed, controlled, and dominated by the more cognitive parts of the psyche." The structure of social relations in education speak to industrial needs. In Dreeben's [12] words, "Although affection is not proscribed in schools, it is expressed less intensely and under more limited circumstances (than in family or peer-group relations). In the long run, matter-of-factness in the accomplishment of tasks governs the relationship between teachers and pupils. . . ." This pattern is repeated in the pattern of rewards in schools. Thus Gebhart and Hoyt [17] find overachievers in high school higher on "Order," and underachievers higher on "Nurturance," and "Affiliation." (Edwards Personal Preference Inventory) Our analysis of the National Merit Scholarship data shows "Originality," "Complexity of Thought," and "Creative Activities" penalized, and similar measures of affective dominance are negatively related to the two main predictors of high grades—CitT and DrA. A similar tension between norms of education and affective, creative development is dramatically illustrated in [18]. Lastly, the Cline study reported above again illustrates that teachers tend to reward the development of cognitive modes but not affective modes, even when these affective modes are conducive to higher levels of cognitive achievement.

Cathection of External Reward. In a situation where the attributes of work and technology are determined essentially independent of human needs and worker control, by criteria of profit and "efficiency"

in the narrow sense, the process of work—as an activity which ideally might provide immediate satisfaction and contribute to individual psychic development as an outlet for creativity, initiative and worker solidarity—naturally acquires little intrinsic subjective value. Moreover, in the absence of a solidary and cohesive social community, and in a situation where workers have essentially no control over the attributes of the product of their work, the internal *goal* of work—the contribution to social dividend—provides no source of gratification and personal reward [20, 22, 23]. The lack of subjective reward of work either in terms of process or goal is the key to what in Marxist terms is called "alienation from process and product" [23], and requires workers to be motivated to conscientious and efficient activity through rewards external to work as such—money or hierarchical status [41].

The development of this motivational capacity is entrusted to socialization mechanisms, among which educational institutions are the most prominent and socially flexible. Indeed, in important respects the system of universal education arose during the Industrial Revolution in response to this need [8, 34]. The structure of social relations in schools reproduce rather faithfully the capitalist work-environment. Learning (the activity) is not undertaken through the student's intrinsic interest in the *process* of learning, nor is he motivated by the *goal* of the educational process (possession of knowledge). Thus the student learns to operate efficiently in an educational environment, unmotivated by either the process or product of his activities—in short, in an alienated educational environment in which rewards are in all cases external: grades, class standing, and the threat of failure. The cathection of such forms of "external reward" is a prime outcome of educational socialization

[12], and doubtless an important contribution to "productive" worker characteristics.

APPENDIX

1. Study: Hansen and Weisbrod (mimeo)
Sample: 2284 predominantly white, non-southern, above-average IQ veterans
Y: Earnings
A: American Air Force Qualification Test
Other Variables: Age, Race
Reduction: 19%
2. Study: Conlisk [10]
Sample: 75 males over a thirty year observation period
Y: Occupational status, scaled by census-average income for the occupation
A: IQ, taken at various ages between ages one and 18.
Other Variables: Parental Income
Reduction: Less than 10%
3. Study: Duncan [13]
Sample: CPS-NORC, Oct. 1964; white men 25-34 years of age CPS-OCG (Oct. Changes in a Generation), March 1962: Non-black, non-farm
Y: 1964 earnings; 1964 Occupational Status
A: early IQ, later IQ
Reduction: between 10% and 25%, depending on the particular Y and A used.
4. Study: Cutwright [11]
Sample: 1% random sample of men registered with Selective Service on April 30, 1953.
Y: Earnings
A: AFQT
Reduction: Between 22% and 35%
5. Study: Duncan, Featherman, and Duncan [15]
Sample: OCG study, all men 20-64 years old; for details see [18] pp. 103 ff.
Y: Status of first job
A: IQ, Army General Classification Test
Reduction: 20%
6. Study: Bajema [1]
Sample: 437 males
Y: Occupational Status, NORC prestige index, age 45
A: Early IQ Terman Group Intelligence, sixth grade
Reduction: 13%
7. Griliches and Mason [26]
Sample: 1964 CPS-NORC veterans file, males, 25-34 years, who have been in army
Y: log actual income
A: AFQT
Other Variables: age, race, sex, SES regional location
Reduction: 12% to 15%, depending on which of the 'other variables' are entered in.
8. Study: Sewell, Haller and Ohlendorf [40]
Sample: a one-third random sample of Wisconsin high school seniors of 1957, follow-up in 1968.
Y: Occupational attainment, using Duncan's (1961) socioeconomic index of occupational status, using data obtained in 1964-65.
A: IQ, Henman-Nelson Test of Mental Ability
E: high school=1; vocational school=1; some college=2; college grad=3
Reduction: 7%
9. Taubman and Wales [46]
Sample: All Minnesota high school graduates of 1936
Y: income in 1953
A: IQ
Reduction: 4%

REFERENCES

1. C. J. Bajema, "Interrelations Among Intellectual Ability, Educational Attainment and Occupational Achievement," *Soc. of Educ.*, 1969.
2. Victoria Bonnell and Michael Reich, *Workers in the American Economy*, New England Free Press, 1969.
3. John Bowers, "Interactive Effects of Creativity and IQ on Ninth Grade Achievement," *Jour. Educ. Meas.*, Fall, 1969.
4. Samuel Bowles, "Towards an Educational Production Function," in *Education In-*

- come, and Human Capital, Studies in Income and Wealth, Vol. 35, National Bureau of Economic Research, 1970.*
5. P. F. Breer and E. Locke, *Task Experience as a Source of Attitudes*, Homewood, Ill., 1965.
6. Frank M. Chambers, "Empathy and Scholastic Success," *Personnel and Guidance Jour.*, Dec., 1957.
7. Victor B. Cline, et. al., "Creativity Tests and Achievement in High School Science," *Jour. Appl. Psych.*, 1963.
8. D. K. Cohen and Marvin Lazerson, "Education and the Industrial Order," paper presented at the AERA meetings, March, 1970.
9. James Coleman et. al., *Equality of Educational Opportunity*, U.S. Dept. of Health, Education and Welfare, Government Printing Office, Washington, D.C., 1966.
10. John Conlisk, "Evidence on the Income-Ability-Education Interrelation," University of California at San Diego, mimeo, 1968.
11. Phillips Cutwright, *Achievement, Military Service, and Earnings*, Social Security Administration, Contract No. SSA 67-2031, May 20, 1969.
12. Robert Dreeben, *On What Is Learned in School*, Reading, Mass., 1968.
13. Otis Dudley Duncan, "Achievement and Ability," *Eugenics Quarterly*, 1968.
14. Otis Dudley Duncan, "Path Analysis: Some Sociological Examples," *American Jour. of Sociology*, July, 1966.
15. Otis Dudley Duncan, David L. Featherman, and Beverly Duncan, "Socioeconomic Background and Occupational Achievement," U.S. Department of Health, Education, and Welfare, Project No. S-0074 (EO-191), May 1968.
16. Sidney A. Fine and Carl A. Heintz, "The Estimate of Worker Trait Characteristics for 4,000 Jobs as Defined in the Dictionary of Occupational Titles," *Personnel and Guidance Jour.*, 1957.
17. G. Gary Gebhart and Donald P. Hoyt, "Personality Needs of Under- and Over-Achieving Freshmen," *Jour. Appl. Psych.*, 1958.
18. J. W. Getzels and P. W. Jackson, "Occupational Choice and Cognitive Functioning," *Jour. of Abnormal and Social Psych.*, 1960.
19. Herbert Gintis, *Alienation and Power: Towards a Radical Welfare Economics*, unpublished Ph.D. dissertation, Harvard University, May, 1969.
20. Herbert Gintis, "Neo-Classical Welfare Economics and Individual Development," Union for Radical Political Economics *Occasional Paper #3*, August, 1970.
21. Herbert Gintis, "Commodity Fetishism and Irrational Production," *Harvard Inst. for Economic Research*, May 1970.
22. Herbert Gintis, "New Working Class and Revolutionary Youth," *Review of Radical Political Economics*, Summer, 1970.
23. Herbert Gintis, "The Structure of Alienation," in Michael Reich et. al. (eds.) *American Society: Conflict and Power*, Prentice-Hall, forthcoming.
24. G. H. Gough, "The Construction of a Personality Scale to Predict Scholastic Achievement," *Jour. Appl. Psych.*, 1953.
25. G. H. Gough and Wallace B. Hall, "Prediction of Performance in Medical School through the CPI," *Jour. Appl. Psych.*, 1964.
26. Zvi Griliches and B. Mason, unpublished data based on on-going research, Harvard University, 1970.
27. J. P. Guilford, "Factors that Aid and Hinder Creativity," *Teachers College Record*, 1962.
28. Eric Hanushek, *The Education of Negroes and Whites*, unpublished Ph.D. dissertation, MIT, August, 1968.
29. Arnold Hauser, "Schools and the Stratification Process," *Am. Jour. of Sociology*, May, 1969.
30. Alfred B. Heilbrun, Jr., "Personality Factors in College Drop-out," *Jour. Appl. Psych.*, 1965.
31. John L. Holland, "Creative and Academic Performance Among Talented Adolescents," in Robert E. Grinder (ed.), *Studies in Adolescence*, New York, 1963.
32. Arthur Jensen, "How Much Can We Boost IQ and Achievement?" *Harvard Educational Review*, Winter, 1969.

33. J. Johnston, *Econometric Methods*, McGraw-Hill, 1963.
34. Michael Katz, *The Irony of Early School Reform*, Harvard University Press, 1968.
35. Kenneth Keniston, *The Uncommitted*, New York, 1960.
36. Cathleen M. Kubiniec, "The Relative Efficiency of Various Dimensions of the Self-Concept in Predicting Academic Achievement," *American Educ. Res. Jour.*, May, 1970.
37. Robert K. Merton, "Bureaucratic Structure and Personality," *Social Forces*, 1940.
38. Joseph L. Norton, "A Study of Worker Traits Requirements for 4,000 Jobs," *Personnel and Guidance Jour.*, 1957.
39. James Scoville, "Education and Training Requirements for Occupations," *Review of Economics and Statistics*, 1966.
40. William H. Sewell, Archibald P. Haller, and George W. Ohlendorf, "The Education and Early Occupational Status Attainment Process," report to the American Sociological Assn., mimeo, San Francisco, Sept., 1969.
41. Herbert Simon, *Administrative Behavior*, New York, 1945.
42. Gene M. Smith, "Usefulness of Peer Ratings of Personality in Educ. Research," *Educ. and Psych. Meas.*, 1967.
43. Gene M. Smith, "Personality Correlates of Academic Performance in Three Dissimilar Populations," Proceedings of the 77th Annual Convention, American Psych. Assn., 1967.
44. Gene M. Smith, "Non-intellective Correlates of Academic Performance," mimeo, 1970.
45. Donald E. Super and John O. Crites, *Appraising Vocational Fitness*, Harper, 1962.
46. Paul Taubman and T. Wales, "Effect of Educational and Mental Ability on Income," Mimeo., University of Penna., Nov. 1969.
47. U.S. Dept. of Labor, Div. of Occ. Analysis, *Dictionary of Occupational Titles*, 2nd. ed., Wash. D.C., GPO, 1949.
48. U.S. Dept. of Labor, Div. of Occ. Analysis, *Estimates of Worker Trait Requirements for 4,000 Jobs*, Wash., D.C., GPO, 1956.
49. Max Weber, "On Bureaucracy," in Hans Gerth and C. Wright Mills, *From Max Weber: Essays in Sociology*, New York, 1958.
50. Max Weber, "The Meaning of Discipline," op. cit.
51. Randall Weiss, "The Effects of Scholastic Achievement on the Earnings of Blacks and Whites," Harvard University honors thesis, March, 1968.

Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data*

By ERIC HANUSHEK
U. S. Air Force Academy

Recent attention to education in the United States represents the merger of concern about efficiency of the educational system and concern about the distribution of educational services, particularly along racial and ethnic lines. However, there is very little guidance on how to satisfy any efficiency or distributional goals through public policy because extremely little is known about the relationship between inputs—particularly inputs available for public policy—and outputs of the educational process. Educational research has been slow in providing definitive answers to public policy questions for several understandable reasons: the subject of the educational process is extremely complex, especially as regards the physiological and psychological aspects; any theoretical development of a learning theory amenable to analysis for policy purposes is absent; and the required data traditionally have not been collected. Previous analyses have yielded some suggestive beginnings, and have provided insights into how the analysis should proceed. This analysis represents a next step of statistical inquiry into the educational process from a

public policy point of view. Three fundamental educational policy questions are addressed: (1) do teachers count? (2) are schools operated efficiently now? (3) what characteristics of teachers and classrooms are important? Past studies have given ambiguous answers to these questions, largely due to inadequate data. Specifically, no data set which supplies accurate historical information on educational inputs at an individual level has been available. This study attempts to provide more conclusive answers by remedying the most glaring data problems for a set of students (third graders) in one school district.

I. Conceptual Model and Data

The major objective of this analysis was to estimate the relationship between variables which can be controlled by public policy and educational output. The basic conceptual model of the educational process which was used can be depicted by Equation 1.

$$(1) \quad A_{it} = f(A_{it^*}, B_i^{(t-t^*)}, P_i^{(t-t^*)}, I_i, S_i^{(t-t^*)})$$

where

A_{it} = vector of educational outputs of the i^{th} student at time t

A_{it^*} = vector of entering achievement levels at time t^*

$B_i^{(t-t^*)}$ = vector of family inputs to education of i^{th} student cumulative from time t^*

* Many people have offered useful suggestions along the way, but special thanks go to Franklin Fisher, John Jackson, John Kain, Herbert Kiesling, Frank Sloan and Finis Welch. This project could not have been undertaken without the help of Wilber Hawkins of the California school district in arranging for data collection and discussing the overall problems. Any remaining errors are, of course, mine. This work was supported by the RAND Corporation, RPN 7502/7510, and The Carnegie Corporation of New York.

$P_i^{(t-t^*)}$ = vector of peer influences of i^{th} student cumulative from time t^*

I_i = vector of innate endowments of i^{th} student

$S_i^{(t-t^*)}$ = vector of school inputs to i^{th} student cumulative from time t^*

This model simply states that educational output (A_{it}), itself a multidimensional factor, is a function of entering achievement (A_{it^*}) and of the influences of the individual's family ($B_i^{(t-t^*)}$), of the influences of his peers ($P_i^{(t-t^*)}$), of his innate abilities (I_i) and of the cumulative school inputs ($S_i^{(t-t^*)}$) over the period being studied.

This abstract model provides a framework for discussion of models of the educational process which can be tested empirically.

Specific variables corresponding to Equation 1 are derived from a combination of past work in the field, theoretical considerations, and sheer data availability. For instance, it is possible to develop many measures of the output of the educational process, such as standardized test scores, juvenile delinquency rates, post-school income streams, occupational choice or level of education completed. Yet, the availability of data has restricted most past studies of education—and this analysis—to examining a single output. This paper concentrates entirely on an analysis of cognitive development as reflected by reading achievement in the third grade. While it is believed that these scores represent differences which are ultimately valued by society, results in terms of this single measure must be considered tentative until there is confirmation in terms of different outputs.¹

¹ There is scattered evidence on the valuation of achievement in W. Lee Hansen, Burton A. Weisbrod and William J. Scanlon, "Schooling and Earnings of Low Achievers," *American Economic Review*, June 1970; Burton A. Weisbrod and Peter Karpoff, "Monetary Returns to College Education, Student Ability and

For public policy purposes, most of the interest in the model centers upon the influence of school characteristics on achievement. However, in order to avoid biased estimates of school effects, the other input vectors must be entered into the statistical analysis. The reasoning behind the inclusion of each of the input vectors is fairly straightforward, and just a brief discussion of specific measures will be given here.² It would be possible to specify a model which did not include an initial achievement level. But, without this input, the data requirements are multiplied considerably since the *entire* past history of all inputs must be known.

Families obviously have considerable impact on education through physical conditions, attitude formation and direct involvement in the educational process. Since these factors tend to be highly correlated with socioeconomic status, of the family, this aspect of education is proxied by father's occupation and family structure. The influence of peers is much the same as that from the family, and, thus, this aspect is proxied by aggregate measures of the socioeconomic status of individuals in a given class or school. While innate abilities are included in the conceptual model, there is no direct measure of this aspect. However, there is reason to believe that biases in the school parameters due to this missing variable are minimal. First, the model with initial achievement measures the "value added" of various inputs and biases will occur only if the missing portion of innate abilities is correlated with the rate of learning (as opposed to the level). Second, at least for whites, it is reasonable to assume that this

College Quality," *The Review of Economics and Statistics*, November 1968; and Randall D. Weiss, "The Effects of Education on the Earnings of Blacks and Whites," *Review of Economics and Statistics*, February 1970.

² A more complete discussion can be found in Eric Hanushek, *The Value of Teachers in Teaching*, RM-6362-CC/RC (Santa Monica: The RAND Corp., 1970).

factor is captured fairly well in the family background variables. This is the case if innate abilities tend to be hereditary and if social mobility is highly correlated with ability. Severe problems, at least in the school portion of the model, do not arise unless there is a mechanism which leads to the correlation of the "nonhereditary" part of innate abilities and specific school resources.

School influences are the focus of this study and will be discussed in more detail than the other inputs. Surprisingly little is actually known about the ways in which schools and teachers affect education, largely as a result of the past fixation on inputs to education rather than outputs. One can impute a set of hypotheses about teacher effects from the behavior of schools in setting pay schedules on teaching experience and educational levels. They must believe that increased experience and further schooling have a positive relationship to educational output.

Other hypotheses can also be found in the actions of school administrators. For example, many persons argue that some forms of student distributions in the schools and classroom (for example, ability tracking or racial and social integration) have a beneficial effect on education. These are testable hypotheses about the relationship between school inputs and achievement. (It must be borne in mind, however, that the tests are restricted to the range of experiences observed.) Further, in recent literature (particularly *Equality of Educational Opportunity*,³ or the Coleman Report) there is a suggestion that one can measure other dimensions of teacher and school quality. These include attitudes of teachers and administrators, verbal facility (and perhaps general ability) of teachers, quality of physical plant, quality

of teacher education, background of teachers, and more.

Of the several studies of the educational process which have been undertaken, two major shortcomings have persisted. First, it has not been possible to match inputs at the individual level, particularly for schools, with the other inputs and outputs of the educational process. Second, there has been a lack of historical data on inputs; most past studies have relied upon cross-sectional data containing only contemporaneous information about inputs. These data problems have introduced considerable doubt into the conclusions of past studies.⁴ A primary objective of this study was to come closer to Equation 1 than had been done previously by eliminating these two sources of data error.

The basic sample of data was drawn from a large school system in California during the summer of 1969. All children in the third grade during the school year 1968-1969 were initially included in the sample. For these 2,445 students, information on family background, scores on the Stanford Achievement Tests, and names of teachers were abstracted from cumulative records. At the same time, all kindergarten through third grade teachers currently in the system were surveyed for information fairly similar to that contained in *Equality of Educational Opportunity*. Information was collected on teacher backgrounds and attitudes, and on specific aspects of schooling. An attempt was made to ascertain their use of time, that is, the division in the classroom between instructional efforts, disciplinary efforts, and administration. Also, a verbal facility test was given each teacher.⁵ The sample used for this

³ James S. Coleman, et al., *Equality of Educational Opportunity* (Washington, D.C.: Government Printing Office, 1966).

⁴ For example, Eric Hanushek and John Kain, "On the Value of *Equality of Educational Opportunity* as a Guide to Public Policy" in Frederick Mosteller and Daniel P. Moynihan (eds.) *On Equality of Educational Opportunity* (New York: Random House [forthcoming]).

⁵ Edgar F. Borgatta and Raymond J. Corsini, *Quick*

analysis was developed by applying two criteria to this group of all third graders. First, individuals were eliminated from the sample if data were not available on both their second and third grade teachers. Second, students were eliminated if both first and third grade achievement test scores were not available. When these criteria were applied, a total of 1,061 students was left in the sample. (A separate analysis of the effects of moving appears to be called for here, but it is beyond the scope of this paper.)

For analytic purposes three different samples were analyzed. As a first step, whites and Mexican-American were separated. (The latter was the only minority group represented in this school system.) There are two reasons for this stratification: (1) the nominal values of the proxies for background inputs do not necessarily have the same meaning for the two groups, and (2) there is no reason to insist on the same model of the educational process for both groups. The ethnic samples were then divided on occupational grounds—fathers in manual or blue collar occupations and nonmanual or white collar occupations. From this, the following three samples were constructed for analysis: white, manual occupation ($n=515$); white, nonmanual occupation ($n=323$); and Mexican-American, manual occupation ($n=140$).⁶

Word Test: Level 2 (New York: Harcourt, Brace and World, Inc., 1964). This test appears to be superior to the test in *Equality of Educational Opportunity* as it appears to give better discrimination among teachers. One complaint voiced about the *EEO* test is that it was too easy. The complete teacher survey can be found in Hanushek, *The Value of Teachers*.

⁶ The decision to stratify will be discussed in terms of statistical tests for sample homogeneity in a later section. These samples are not exhaustive. Children with only mothers or no occupation reported for fathers were not included. For whites, these groups totaled 36 students; for Mexican-Americans, these groups plus the nonmanual occupation group totaled 47. These samples were too small to study separately, and, thus, they were ignored.

II. Do Teachers Count?

Recently there has been considerable controversy among those analyzing education as to whether teachers count in the educational process. This arises from interpretation of past empirical work, namely the Coleman Report. However, since our sample experience did not include children without teachers, the only testable hypothesis is whether or not there are differences in teachers that lead to differences in achievement among students. In other words, does it matter which teacher a student has, or are all teachers perfectly substitutable?

This test is done by constructing a series of dichotomous variables, T_{ij} , for each teacher in the sample. If the j^{th} student has the i^{th} teacher, $T_{ij}=1$ for him and $T_{kj}=0$, where $K \neq i$. The complete model looks like:

$$(2) \quad A_{j3} = \sum_{i=1}^n t_i T_{ij} + aS_j + bA_{j2} = u_j$$

where

A_{j3} = achievement in 3rd grade of the j^{th} student

$S_j = 1$ if j^{th} student is female; $= 0$ otherwise

A_{j2} = achievement in grade 2 of j^{th} student

(Other analysis of these data indicates that sample stratification by ethnic and occupational background is an adequate way to allow for family inputs. Thus, no explicit family background measure was included here.) In this formulation it is possible to ask whether the individual classroom coefficients are significantly different from a constant, or whether there are any differences among teachers in the sample in terms of their contribution to achievement gains.

At least two students from a sample had to be in a class with a teacher before the

TABLE 1—*F*-STATISTICS FOR NULL HYPOTHESIS OF UNIFORM TEACHER EFFECTS

Sample	<i>F</i>	d.f. ^a	<i>R</i> ²
<i>3rd Grade</i>			
White, manual	2.03 ^b	(69, 426)	.71
White, nonmanual	1.57 ^b	(57, 247)	.76
Mexican-American, manual	.78 ^c	(29, 78)	.68
<i>2nd Grade</i>			
White, manual	2.96 ^b	(55, 440)	.68
White, nonmanual	2.39 ^b	(48, 264)	.71
Mexican-American, manual	1.09 ^c	(26, 82)	.64

^a Degrees of freedom.

^b Statistically significant at the .01 level.

^c Not statistically significant at the .10 level.

student and teacher were included in the analysis. For the three samples tests for differences in the t_i 's in both the third and second grade were performed. (For the second grade analysis, the dependent variable is A_{j2} and one exogenous variable is A_{j1} .) The results of these six *F* tests for equality of coefficients are depicted in Table 1. For whites, the hypothesis of no teacher differences is rejected at the 1 percent level. However, for Mexican-Americans it is not possible to reject the hypothesis of no teacher differences at the 10 percent level. In other words, the teacher appears to count for whites of all social strata but not for Mexican-Americans.

One qualification is needed before any further interpretations are made. Since these students had only one teacher during the year, it is impossible at this stage of the analysis to distinguish between the effects of particular teachers and a classroom composition effect. There is no independent observation here for a given teacher with several different classrooms. This problem will be dealt with directly in the next section.

This analysis suggests that the Mexican-Americans at this lower grade level are not getting much out of school. On the average they tend to progress at a rate of about one-half grade level per year or 50 percent

of the national average for reading achievement gains, regardless of which teacher they have. It is possible that the classroom composition exactly offsets teacher differences or that teachers are matched with Mexican-American classes to equalize gains. However, this seems highly unlikely since the matching needed to achieve no teacher-classroom effects calls for putting the best teacher-classroom combination with the room of worst "gainers," etc. Moreover, since the white children are sensitive to teacher-classroom differences, as indicated by Table 1, a finding of no differences for Mexican-Americans—when in fact white differences exist—implies that teachers are distributed only in conjunction to the Mexican-Americans in the class. Yet, the proportion of Mexican-American ranges from 6 percent to 63 percent in the 30 third grade classrooms that have more than 2 Mexican-American students.

III. Characteristics of Teachers

The preceding section suggested that the performance of white students is dependent upon the specific teacher and classroom associated with the student. For policy purposes it would be useful to identify the characteristics that contribute to increased performance. This phase of the analysis was accomplished by introducing a variety of quantitative teacher and classroom characteristics into an overall model of student achievement. This was done for both the white manual occupation and white nonmanual occupation samples.

The first step was to estimate models which included only the "pay parameters" of teacher experience and hours of graduate education to measure schools. After allowing for differences in entering achievement, family background, and school composition, these variables for second and third grade teachers were always statistically insignificant; all *t*-statistics were less

TABLE 2—VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS

Variable	White, Manual		White, Nonmanual		Definition
	Mean	Std Dev	Mean	Std Dev	
A_3	55.74	19.1	64.82	16.8	Stanford Achievement Test raw score—3rd grade
F	.50	.5			Sex: =1 for female =0 for male
R	.08	.3			Repeat grade: =1 if a grade was repeated; =0 otherwise
A_1	35.17	15.1	42.43	15.8	Stanford Achievement Test raw score—1st grade
D	17.93	18.8			% of time spent on discipline by 3rd grade teacher
T_3	66.90	15.8			Quick Word Test score—3rd grade teacher
Y_3	1.91	1.6	2.02	1.7	Years since most recent educational experience—3rd grade teacher
T_2	68.41	19.0			Quick Word Test score—2nd grade teacher
Y_2	2.64	2.6	1.88	1.7	Years since most recent educational experience—2nd grade teacher
C			.19	.4	Clerical occupation; =1 if father in clerical job; =0 otherwise
S_3			7.85	8.1	Years of experience with this socioeconomic level—3rd grade teacher
S_2			7.94	6.1	Years of experience with this socioeconomic level—2nd grade teacher

than 1.25. This indicates that schools are now operating inefficiently but gives minimal guidance to school administrators. The remainder of the paper attempts to identify the aspects of schools and teachers which are important in education.

The estimates for the white manual sample are displayed in Equation 3. Variable definitions, means, and standard deviations appear in Table 2.

$$(3) \quad A_{j3} = 20.8 + 2.81F - 6.38R$$

(2.3) (−2.8)

$$+ .79A_{j1} - .07D$$

(18.8) (−2.1)

$$+ .09T_3 - .57Y_3$$

(2.4) (−1.5)

$$+ .06T_2 - .68Y_2$$

(1.9) (−2.9)

$$R^2 = .51 \quad SE = 13.5$$

(*t*-statistics are displayed below each coefficient; *SE* is the standard error of estimate.)

This model presents an interesting view of teachers. The teacher characteristics that appear to be important are not the characteristics that are purchased by schools. For both the second and third grade teachers, the score on the verbal facility test (*T*) and the recentness of education (*Y*) are the most important factors.⁷ Additionally, there is a "quasi teacher" characteristic of percent of time spent on discipline by the third grade teacher. Each of these has important implications for school operations.

The verbal facility test (*T*) probably

⁷ The model was not constrained to have the same characteristics for second and third grade teachers; this results from the analysis of various characteristics without constraint. If we test the joint hypotheses that all four strictly teacher characteristics together have no effect on education, we reject at the .01 level with $F_{4,500} = 5.68$.

TABLE 3—SIMPLE CORRELATIONS FOR TEACHER CHARACTERISTICS—WHITE MANUAL SAMPLE

	<i>D</i>	<i>T</i> ₃	<i>Y</i> ₃	<i>T</i> ₂	<i>Y</i> ₂	Exper ₃	Units ₃	Exper ₂	Units ₂
<i>D</i>	1.00								
<i>T</i> ₃	-.19	1.00							
<i>Y</i> ₃	.01	.08	1.00						
<i>T</i> ₂	.07	.19	.13	1.00					
<i>Y</i> ₂	-.14	-.09	.11	-.19	1.00				
Exper ₃	-.14	.37	.11	.17	.05	1.00			
Units ₃	-.09	.01	-.14	.09	.12	.53	1.00		
Exper ₂	.09	-.11	-.01	-.09	.20	-.18	-.11	1.00	
Units ₂	-.02	.07	.01	-.03	-.15	-.02	.03	.43	1.00

plays two roles: first, it is a measure of communicative ability; second, as the authors of the test point out, it can be taken as a quick measure of overall intelligence or general ability. Thus, general ability seems important, regardless of formal training. There are some important policy implications surrounding the verbal test measure of teacher quality. By interchanging teachers at the top and bottom of the verbal ability scale for this system, achievement changes by .2 to .4 grade levels.⁸ This is very significant given the powerful effect of a student's early education on later achievement as evidenced both by the increasing grade level disparities in the Coleman Report and by the strength of first grade achievement in Equation 3. Further, since this test has national norms, it is possible to get some idea of how the teachers being hired in this system rate alongside other college graduates. The mean score of 68 places the teachers in this sample slightly under the median for female college graduates. Thus, this system is not being successful in attracting the best people.

In addition to teacher ability as measured by the verbal facility test, the recentness of educational experiences (*Y*) has a significant effect on educating students.

⁸ This is calculated by changing only the third grade teacher verbal score for the lower limit and both second and third for the upper limit. The scores are changed from 40 to 96 to represent the range found in the data. (Maximum score is 100.) The resulting achievement score is then converted to grade level equivalents.

This seems to provide the rationale for encouraging or requiring teachers to return to school periodically. However, as indicated by the results of analyzing graduate units and the effects of Master's degrees, it does not really matter whether the teacher is enrolled in an advanced degree program or is taking many courses. Education of the second and third grade teacher in the last year as opposed to five years ago would be worth .2 to .3 years of reading achievement to a given third grader.

Finally, there is the measure of discipline time (*D*) that was labeled as a quasi-teacher characteristic. Certainly, an interaction between the classroom and the teacher is reflected in this variable. However, as expected, the more time spent on disciplinary matters, the lower the achievement level of the class. It does suggest that efforts to reduce such time could be beneficial. These efforts would include using principals or assistant principals or even teacher's helpers as disciplinarians.

It is immediately obvious that these are not the characteristics of teachers that are currently being purchased. Certainly if there is an excess supply of teachers, schools can be selective in hiring and can attempt to evaluate the general ability of teachers. However, casual observation suggests that the most selective (suburban) systems weight previous teaching experience heavily. Moreover, as suggested by the simple correlation matrix for teacher characteristics displayed in Table 3, the

purchased factors for the third and second grade teachers (experience and units of graduate work) are not highly correlated with the characteristics included in the model.

A model for the white nonmanual population was also estimated.⁹ The results of this analysis, shown in Equation 4, provided a different set of teacher characteristics that seemed important. (Variable definitions, means, and standard deviations are displayed in Table 2.)

$$\begin{aligned}
 (4) \quad A_3 = & 35.0 + .72A_1 - 5.1C \\
 & \quad (16.0) \quad (-3.0) \\
 & - .79Y_3 + .10S_3 \\
 & \quad (-1.9) \quad (1.2) \\
 & - .66Y_2 + .20S_2 \\
 & \quad (-1.7) \quad (1.8) \\
 R^2 = & .52 \quad SE = 11.8
 \end{aligned}$$

The characteristics which seem important for the white children from nonmanual occupation families are not entirely the same as for children from families with manual occupations. The recentness of education is again a significant factor, with approximately the same effect here as in the previous model. However, while teacher verbal ability does not appear to

be significant, experience with this socioeconomic group assumes importance. As would be expected, the correlation between total experience and socioeconomic group experience is quite high. The simple correlation for third grade teachers is .8. Thus, the present policies of paying for experience could be reasonable in this case. Recentness of education is, however, only slightly correlated with the pay factors, as in the manual occupation sample.

In comparison to the manual occupation sample, the coefficient estimates are not as reliable in the nonmanual sample. Although there is a smaller standard error of estimate for the nonmanual model, this is coupled with a smaller variance in overall achievement; the R^2 's in the two models are almost equal. The estimated coefficients in Equation 4 do indicate that teachers have less effect on these nonmanual children. Although the effects of recent education are roughly the same in the two models, the potential for change in achievement through increasing verbal facility, or general ability, is considerably greater. Not only can verbal facility be changed more rapidly—since experience usually comes by aging rather than hiring—but given percentage changes in verbal quality have a considerably larger effect on student achievement.

The previous discussions of Equations 3 and 4 must be taken within the context of the overall model. In the process of developing these models, several other hypotheses about educational inputs were tested and rejected. In particular, variables measuring school composition in terms of occupational distribution, ethnic distribution, and achievement distribution; variables measuring objective background characteristics of the teachers such as socioeconomic background, college major, and membership in professional organizations; and variables measuring subjective factors such as attitudes toward types of students were tested and found to

⁹ A word on sample stratification is necessary. A formal test for equality of coefficients between the two white models was performed. The methodology of this test can be found in Franklin M. Fisher, "Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note," *Econometrica*, March 1970. When testing the entire model and restricting the models for both samples to the form of Equation 3, the hypothesis of coefficient equality was rejected at the .025 level ($F_{9,820}=2.13$). However, since the principal interest centers upon the teacher characteristics, a test of just this subset of variables seems more appropriate. When this test is performed, the results are inconclusive since $F_{4,820}=1.57$ when the critical value for those degrees of freedom at the .10 level is 1.94. Thus, it is not possible to reject the hypothesis of homogeneity with a high level of confidence; yet, at the same time the evidence does not seem strong enough to pool the sample. Since both samples are quite large, the loss in efficiency by not pooling would not be large, and the decision was made to stratify.

have statistically insignificant effects on the students' achievement. The implication that arises from both these "nonresults" and from the models presented is that we do not have very good measures of teachers. We can identify a few objective factors that appear to affect education. Yet most of our notions about important attributes of teachers are probably too simple. Although teachers do appear important in the model, precise decision rules for hiring teachers are not readily available at this time. It appears that objective characteristics of teachers and classrooms are insufficient and that it is necessary to measure behavior in the classroom better.

IV. *Summary and Conclusions*

From this study three conclusions are apparent. It should be borne in mind that these conclusions derive from a sample of one school system; therefore, some caution should be used in generalizing to other systems. First, the present set of hiring practices leads to an inefficient allocation of resources. The analysis indicates that teaching experience and graduate education do not contribute to gains in student achievement scores. Moreover, the characteristics that do matter are not highly correlated with these factors. Yet these attributes are being purchased by the school district. Since turnover is costly,

some average experience level over one year would be reasonable. However, the current average of over 11 years is certainly excessive. Second, teachers do not appear to count for Mexican-American students in the sense that different teachers and different classroom compositions do not affect the achievement outcome of the Mexican-American students. This might well be a language problem, but there is no direct measure of this. Third, the attempts to provide a set of measurable characteristics which schools could hire and control to affect achievement did not produce clear-cut answers. There is a considerable part of teaching that cannot be explained by a set of fairly standard variables measuring teachers and classrooms, particularly for whites in white collar families.

It would be imprudent to generalize too much from these findings. They refer to one school system, one measure of output, and one particular grade level in elementary schools. For this reason this study is best looked upon as being suggestive rather than definitive; as being a prototype rather than a final analysis. Yet the evidence is beginning to mount and seems to indicate fairly consistently that the past ways of operating schools leave something to be desired.

Ability and Schooling as Determinants of Lifetime Earnings or If You're So Smart, Why Aren't You Rich?

By JOHN C. HAUSE

National Bureau of Economic Research and University of Minnesota

"Ability: the natural equipment to accomplish some small part of the meaner ambitions distinguishing able men from dead ones." *Ambrose Bierce*

The role of ability in determining earnings either directly or indirectly by increasing the earnings productivity of education is currently not well understood. Even Bierce's broad definition of ability is inadequate for empirical economic study. Whether we are concerned with highly specific ability (e.g., the admirable and rare capacity of a person to extract all the rent in salary or consulting negotiations) or with general ability (e.g., as measured by a single or by multiple test scores), ability is presumably the outcome of genetic, environmental, and experience factors that are difficult to disentangle. Fortunately, for some problems it may be unnecessary to carry out a decomposition

of ability as the following simple model illustrates. Suppose that the earnings expected by an individual can be expressed as a function of measured (general) ability at t_0 , a collection of personal characteristics at t_0 , additional schooling (or other investment in earnings capacity), and time for $t \geq t_0$:

$$(1) \quad E(t) = f(A_0, P_0, S, t).$$

A_0 and P_0 specify the initial state of the person at t_0 . Although P_0 and A_0 may represent simple identifiable factors (such as sex or race) or complicated functionals of factors operating before t_0 , it may be that A_0 and P_0 summarize adequately the initial state for analyzing how S changes the earnings profile $E(t)$. If this condition holds, so that a complete biography of the individual from conception until t_0 can be avoided, a value added analysis of the schooling increment is feasible.

In most empirical studies A_0 is simply a test score (or set of scores) that partially characterizes the initial state of the person. However, we usually have some loose hypotheses about how different values of A_0 alter $E(t)$ for given levels of schooling after t_0 (including $S=0$). The first section discusses some of these hypotheses, while the second considers several cohorts of individuals for which data are available on ability, achievement, other personal and background characteristics, and earnings for two or more points in time.

* University of Minnesota and National Bureau of Economic Research. The empirical portion of this study would have been impossible without the data carefully assembled by D. C. Rogers. I am indebted to him for allowing me to use his basic data tape. I am similarly indebted to Professor Torsten Husén for allowing me to use the data on Swedish males and to Ingemar Fägerlind for the Swedish earnings data. The American Institutes of Research prepared the Project Talent data tape. Comments by Finis Welch on an initial draft of this paper were characteristically perceptive and useful. Computations were effectively carried out by J. Sanguinetti, M. Sternfeld, and A. L. Norman. Margareta Forselius and Karlis Goppers gave much assistance with the Swedish data. Much of the research was carried out with the support of National Science Foundation Grant GS-1797.

I. *Some Hypotheses on Ability, Schooling, the Earnings Profile, and Lifetime Earnings*

"Ability" is usually defined as the power to do something. Many of the tests designed to measure ability have been developed in an educational context in which the relevant power is the power to learn. Learning through formal schooling and learning those things that increase a person's economic productivity are not simple or identical, but there is clearly some overlap. Thus one expects there is a positive, but imperfect correlation between abilities to learn from these two sources of knowledge.

To my knowledge, neither learning theory nor economic theory have been developed to the point where a powerful theory of the earnings profile can emerge. Even so, interpreting ability loosely as "learning power" immediately suggests several simple hypotheses. Consider the earnings profiles of a cohort of people with the same level of schooling completed. If people with more ability learn the same job skills as others but with greater speed, their earnings profiles will rise more rapidly than those with less ability as long as their economic productivity is increasing more rapidly. If full job competence is attainable by both people of lower and higher ability, the influence of ability disappears after a

period long enough to allow those with less ability to attain full productivity. On the other hand, differential ability may limit the complexity of skills that people are able to master. In this case, an ability effect on earnings can persist over time as long as more complex skills yield higher earnings.

When ability is measured by a test of capacity to learn from formal schooling, there is a further reason for expecting a positive interaction of schooling level and measured ability on earnings. An empirically significant positive association exists both between level of schooling attainment and earnings (e.g., Table 2) and between ability and level of schooling attainment (e.g., Tables 1 and 2). Research has not yet clearly established what components of schooling are the most important determinants of earnings differentials at different levels of schooling. A number of schooling output components have been suggested, including (1) specific skills to perform well-defined tasks, (2) general cognitive skills that enable people to locate and handle information more efficiently, (3) "social skills" that increase the capacity to deal with others, (4) development of greater rational foresight and self-discipline, and (5) accustoming people to certain attitudes (e.g., to obeying orders) and to routine tasks that increase their

TABLE 1—PERCENTAGE OF MALE HIGH SCHOOL GRADUATE QUESTIONNAIRE RESPONDENTS *NOT* ENTERING COLLEGE WITHIN ONE YEAR OF GRADUATION

Aptitude Percentile	Family's Finances				
	Extremely Wealthy or Wealthy	Well-to-Do	Comfortable	Have Necessities	Barely Make a Living
90-99.9	3.7	5.1	7.7	14.5	16.1
75-89.9	13.5	10.4	19.5	27.0	21.7
50-74.9	25.3	23.0	40.2	41.8	51.2
25-49.9	39.4	44.2	59.1	70.3	78.9
0-24.9	83.5	72.5	77.0	81.6	87.6

Source: Project Talent, *The American High-School Student* (Tables 11-18), Final Report for Coop. Project 635, U.S. Office of Education, 1964.

TABLE 2—MEANS AND STANDARD DEVIATIONS* OF LOG EARNINGS, IQ, AND BACKGROUND VARIABLES BY EDUCATION LEVEL (ROGERS)

Edu- cation Level	Sample Size <i>N</i>	LE65	LE60	LE55	LE50	LDE4%	IQ	SCE	SCL	RC	RJ	PS	NM
<i>E</i> ₁	60	8.857 (.326)	8.708 (.281)	8.664 (.279)	8.569 (.335)	11.836 (.221)	95.9 (11.8)	.033	.917	.750	—	.033	.050
<i>E</i> ₂	117	9.001 (.392)	8.764 (.336)	8.662 (.320)	8.550 (.338)	11.872 (.324)	102.3 (11.1)	.034	.829	.650	.017	.094	.077
<i>E</i> ₃	51	9.262 (.557)	9.057 (.439)	8.900 (.428)	8.668 (.478)	12.070 (.429)	107.8 (9.59)	.196	.529	.431	.039	.235	.039
<i>E</i> ₄	70	9.434 (.599)	9.244 (.481)	9.004 (.429)	8.692 (.528)	12.204 (.521)	115.6 (11.0)	.429	.414	.314	.043	.471	.067
<i>E</i> ₅	47	9.640 (.574)	9.414 (.607)	9.061 (.502)	8.525 (.624)	12.262 (.445)	117.3 (10.0)	.446	.319	.298	.170	.383	.042

* Numbers in parentheses are standard deviations. Source—data tape of D. C. Rogers. Sample is based primarily on male eight graders (most from Connecticut) who were given IQ tests in 1935. Earnings are based on questionnaire responses in 1966. Earnings from 1965 are reported annual earnings plus imputed self-employment income from personal labor services. Earnings for 1960, 1955, and 1950 are full-time annual earnings based on reported monthly or weekly salary. Lifetime discounted earnings were obtained by Rogers through an involved smoothing and extrapolation procedure.

productivity in modern economic organizations. It seems plausible that measured ability is primarily associated with general cognitive skills and with the capacity to acquire some of the more complex specific mental skills. Since these skills (and skill-levels) depend on both schooling and measured ability, one expects a positive interaction of schooling and ability on earnings unless these skills do not affect earnings.

The relative slope of the logarithm of the earnings profile for high and low ability people is not obvious. *A priori*, initial earnings of people first entering the labor force could have a positive, zero, or negative simple correlation with ability. A positive correlation could indicate that those with higher ability are initially more productive and that employers can determine this fact at the time of hiring. In this case, there is no guarantee that the percentage rate at which these people acquire the relevant job skills is greater than

the rate for less skilled people. A low positive or zero simple correlation between ability and initial earnings could reflect imprecise information by employers about the current and potential productivity of new entrants to the labor force at the time they are hired. The reassessment of employee productivity gained through experience and the higher speed with which the more able workers acquire job-relevant skills should combine in this case so that the percentage rate of increase in earnings is initially higher for those with more ability. A negative simple correlation between initial earnings and ability could arise if ability is a strong complement of on-the-job training which may be paid for by reduced initial earnings. In this case, at some point the relative earnings of high ability people would have to rise more rapidly than those of people with less ability to make worthwhile the investment financed by reduced levels of initial earnings.

Even for a cohort of people with same level of schooling and ability, the growth of individual productivity over time may differ between jobs. In the absence of non-pecuniary occupational tastes, there would be a tendency for entry rates to different occupations to be governed by the condition that they have the same (net) present value of earnings. Thus significant dispersion of earnings at different points along the earnings profile may be compatible with relatively little dispersion in the net present values. Since people presumably take into account prospective profiles of earnings over time and not merely the earnings for a single year when they make decisions about schooling and occupation, the attempt to establish statistically the determinants of the present value of earnings (or closely related functions of the earnings profile) plays an important part for understanding these decisions. The preceding remarks on ability and how it may affect the earnings profile also imply that there is likely to be a positive correlation between ability and discounted earnings.

We consider next the relationship between ability and discounted earnings for different levels of schooling. A number of expository and statistical models have presented earnings (usually for a single year after earnings profiles have flattened out) as a linear function of schooling, ability, and an uncorrelated random variable. (See, e.g., W. Lee Hansen, Burton A. Weisbrod, and William Scanlon or Orley Ashenfelter and Joseph D. Mooney.)

$$(2) \quad Y = \beta_0 + \beta_1 S + \beta_2 A + u.$$

This relationship does not seem plausible. It implies that schooling and ability are perfect substitutes (because of the linear form). More important, it implies that the marginal product of additional schooling is independent of ability (because of separability). The latter assumption is implausible because of two well-known facts.

First, there is a systematic tendency for higher ability people (measured by IQ or other tests) to acquire more schooling than others. Second, the opportunity cost of foregone earnings is a large part (more than half) of the cost of obtaining higher levels of schooling for most people (e.g., see Theodore W. Schultz). Equation (2) implies that the opportunity cost of acquiring additional schooling is greater for more able people, yet this schooling yields the same increment to earnings to people independently of ability. Thus the economic incentive to acquire additional schooling implied by this model is greater for those with less ability, and their expected rate of return would be higher.

An alternative specification that captures the opportunity cost of acquiring schooling in a more plausible way replaces the level of earnings Y by $\log(Y)$ in equation (2). In this case, the level of earnings for people with different ability increases equiproportionally with schooling. Even this specification provides no economic reason why there is such a strong tendency for higher proportions of people with greater ability to acquire more formal education. Table 1 (which contains some response bias) shows the strength of this relationship for recent male college entrants. Although there is a greater likelihood for sons of high income families to enter college at all ability levels, the ability-college-entrance relationship is very strong within family income classes.

A simple model in which education is acquired solely for earnings capacity, and in which perfect foresight and a perfect capital market for funds to support schooling are assumed, implies an equilibrium when the flows of different ability people to different terminal levels of schooling leads to relative wages such that net present values of earnings are the same for people with the same ability but different

schooling attainment. In such a world, within schooling class regressions of the logarithm of earnings on ability would result in the coefficient of ability being roughly the same for different schooling levels. However, the very imperfect market for educational loans (and perhaps uncertainty) might well lead to a tendency for the coefficients of ability in these semi-logarithmic regressions to rise with education. The hypotheses of this section are examined empirically in Part II.

II. Empirical Results

Although four samples of cohort data are being examined to study the ability-schooling-earnings relationship, extensive calculations have been completed so far only on the important but small sample of New England males (primarily from Connecticut) obtained and previously reported on by Daniel C. Rogers. Rogers' data are handled somewhat differently in this paper and these differences lead to several modest modifications of his results and to some complementary conclusions. Preliminary results from other data sources are mentioned briefly at the end of this section.

Table 2 shows the means of the earnings, ability, and background variables (and standard deviations of nondichotomous variables) by level of schooling from the Rogers sample. The five schooling levels are E_1 high school without graduation, E_2 high school with graduation, E_3 some college, E_4 one college degree (and perhaps additional study, and E_5 two or more college degrees. Most sample members were eighth graders in 1935, and 85 percent are veterans of World War II. These features of the sample make the E_3 group atypically heterogeneous. E_3 includes men who started college shortly before the war, entered the military, and didn't return to college. It also includes those college dropouts who initially entered college after completing military service, attracted in

part by the relatively low costs of college attendance because of GI bill subsidization. The motivations leading men to enter college but not graduate are unusually unclear and heterogeneous in this sample, and one should interpret with caution particular results for this subgroup.

The background variables used in this study are subpopulation dummy variables for social class ($SCH=1$ for the highest of four social classes, $SCL=1$ for the lowest), religion ($RC=1$ for Catholic background, $RJ=1$ for Jewish background), private school attendance ($PS=1$ for precollege private schooling), and marital status ($NM=1$ if not married in 1965).

The earnings means all increase with schooling except for 1950 when the E_6 group has very little postschool job experience and several trivial reversals in other classes. The standard deviations of the log of earnings are lower for E_1 and E_2 levels, which suggests in principle that weighted regression should be used if all classes are pooled in one regression. The table shows the positive association of IQ and schooling level, a relationship that suggests the strong possibility that schooling and IQ have a positive interaction on earnings, a hypothesis proposed in Section I. The positive association of high socioeconomic class and schooling level is consistent with the similar association in Table 1. Marital status is not systematically related to schooling attainment, but is included in the empirical work because of the strong tendency for unmarried men to have lower earnings than married males. The entire set of background variables is frequently used in regressions reported in this paper and is denoted by X^* in the following discussion and tables.

The schooling subgroups in this sample are unpleasantly small, which leads to large standard errors in many of the parameter estimates. Even so, there are some suggestive patterns that broadly con-

form to some of the hypotheses in Section I although they cannot be confirmed with high statistical significance.

Several theoretical and empirical arguments in the preceding section explain why schooling and ability are unlikely to be perfect substitutes in producing earnings. Table 3 provides some evidence on this point by showing the linear regressions of 1965 earnings and discounted lifetime (at 4%) on IQ and the background variables X^* . Only the IQ coefficient is reported. The pattern of IQ coefficients (except in E_3) is broadly consistent with the belief that the coefficient of ability increases with education level in linear regressions. IQ is only trivially related to earnings for the lowest schooling level, but appears to make a moderate empirical difference in earnings as the schooling level rises. An approximate chi-square test of the statistical hypothesis that the 1965 IQ coefficients are equal across education classes (excluding the peculiar E_3 class) indicates the probability of the null hypothesis being less than .1.¹ This result and the array of IQ coefficients in Table 3 suggests that there is positive interaction between IQ and education level and that the linear model is misspecified by not allowing for the interaction.

The coefficient for the E_3 group (some college) is anomalous. This result may be sample-specific for historical reasons already mentioned or there may be some unobservable factor that leads to self-selection by those who terminate their school at this level. Still it is unclear why the effect of IQ on earnings should be eliminated.

Table 4 contains some regressions of the

TABLE 3—LINEAR REGRESSIONS OF EARNINGS ON IQ

Education Level	Sample Size N	1965 Earnings		Discounted Lifetime Earnings (4%)	
		β_{IQ}	R^2	β_{IQ}	R^2
E_1	60	-3.5 (26.9)	.16	-115 (367)	.19
E_2	117	74.6 (35.8)	.19	756 (592)	.09
E_3	51	-2.2 (127.2)	.34	-589 (1606)	.23
E_4	70	132.5 (88.7)	.31	1670 (2395)	.23
E_5	47	223.0 (154.5)	.19	1968 (1754)	.19

logarithm of earnings on IQ and other variables to observe whether IQ differentials affect earnings at all educational levels equiproportionally. As before, IQ continues to have a very weak association with earnings of high school nongraduates. Aside from the continued anomaly of the small coefficient for E_3 , the pattern of IQ coefficients in 1965 and for discounted lifetime earnings across schooling levels is consistent with the argument in Section I that ability should tend to increase earnings at least proportionally for increasing levels of schooling. Indeed, the IQ coefficient on 1965 earnings and discounted earnings appears to jump substantially for the highest education level, a suggestive result, although in this small sample the difference in the E_4 and E_5 coefficients is not statistically significant.

In order to study how IQ affects earnings for a given schooling level, regressions can be run using the difference in the logarithm of earnings for distinct years as dependent variables. The coefficient of an independent variable can then be interpreted as the percentage change in earnings over time associated with the variable. The coefficients of IQ that would be ob-

¹ For this test, a weighted mean was constructed from the IQ coefficients of the 1965 regressions in Table 3. Let $w_i = (1/\sigma_{\beta_i}) / \sum_j (1/\sigma_{\beta_j})$ where σ_{β_i} is the estimated standard error of β_i , the IQ coefficient of schooling group E_i . Let the weighted mean $\beta_w = \sum w_i \beta_i$. Then $\sum [(\beta_i - \beta_w) / \sigma_{\beta_i}]^2$ is approximately chi-square with three degrees of freedom (for the four education classes).

TABLE 4—COEFFICIENTS* ON IQ FROM REGRESSIONS OF
LOG-EARNINGS ON IQ AND OTHER VARIABLES
(Year or Discount Rate)

Education Level	N	1965	1960	1955	1950	4%	Other Variables
E_1	60	.024 (.35)	.14 (.29)	.20 (.29)	-.27 (.36)	-.01 (.24)	X^*
E_2	117	.70 (.32)	.36 (.29)	.32 (.28)	.28 (.30)	.45 (.27)	X^*
E_3	51	.36 (.78)	.17 (.66)	.32 (.62)	.04 (.68)	.00 (.63)	X^*
E_4	66	.91 (.62)	.69 (.54)	.52 (.50)	.74 (.70)	.42 (.60)	X^*
E_5	47	1.32 (.90)	1.01 (.99)	-.18 (.97)	-.17 (1.11)	.78 (.75)	X^*
Pooled	345	.49 (.23)	.46 (.20)	.29 (.20)	.15 (.22)	.35 (.19)	X^* and education level dummies

* Numbers in parentheses are standard errors. Standard errors and coefficients have been multiplied by 100. The 1965 IQ coefficient for E_5 , for example, implies a 1.32% rise in earnings with a one-point rise in IQ.

NOTES—The coefficients of determination in these regressions in 1965 vary from .17 to .35 (pooled class is .38) from .28 to .03 for 1950, and from .36 to .11 for the 4% discount rate. The social class variables are low in statistical power and much smaller in effect than the religious, marital status (1965) and private school dummies.

tained from this calculation can be read directly from Table 3 by taking the difference in the coefficients for any pair of years. For E_5 (two or more college degrees), IQ appears to have an effect that increases substantially over time. Although occupational information was unavailable, this result might be due in part to substantial earnings by high ability professional men whose earnings may increase rapidly after their lengthy training is completed. (The initial small negative correlations of IQ with earnings for this group in 1950 and 1955 may also be due to the late labor market entry of these professionals.) For most schooling levels there appears to be a tendency for ability to become more important as labor force experience increases. A puzzling exception is the cohort of college graduates with one degree. The IQ coefficient is singularly large for the first reported year, 1950. The general pattern of Table 4, except for the E_3 (some college) group, suggests a tendency for the relative influence of ability

to increase as one moves along the lifetime earnings profile.

On the basis of this sample, what conclusions can be drawn about the important problem of bias in the returns from education if ability is ignored? It seems to be well established that mean IQ increases the higher the school level. If IQ is positively correlated with earnings at each level, then it is clear that the apparent increase in earnings by taking the difference in the mean earnings at each level exaggerates the potential gain for a person of given ability in making the transition from lower to higher level. For the Rogers sample, Table 2 indicates that mean IQ of college graduates is about 13 points beyond high school graduates; the corresponding differential in IQ for college graduates with two degrees over high school graduates is 15. The coefficients on IQ for levels E_4 and E_5 in Table 4 multiplied by the corresponding differentials imply that lifetime discounted earnings (at 4%) of the mean high school graduate

would be about 6% and 8% less than the corresponding means of those who attained the E_4 and E_5 levels. A similar calculation based on the 1965 earnings IQ coefficients implies that the mean high school graduate (who terminated his education with high school) would earn 8% and 18% less than the mean E_4 and E_5 individuals, respectively. Thus, this sample indicates an empirically significant bias exists that overstates the apparent increase in earnings for high school graduates who subsequently take one or more college degrees. Because the IQ-earnings relationship is negligible for high school nongraduates, there is no overstatement of the increased earnings obtained by those who actually completed high school (although there would be an overstatement of the gains to a person whose ability is that of the mean high school nongraduate).

The sample provides a modest but positive rationalization of the strong association between schooling attainment and IQ. The discounted lifetime earnings coefficient for those with two or more degrees is larger in magnitude than the high school coefficient. The product of the difference in the coefficients and the difference in the mean IQs indicates that lifetime discounted earnings increase 5.4% more for a person with IQ 117 than for a person with IQ 102, if they both have two degrees.²

² This magnitude of interaction between IQ and schooling is considerably larger than the 1% increase obtained by Rogers (Table 9, p. 115). The difference arises in part because Rogers uses an age-in-school variable that is correlated with IQ and because of other differences in formulating a statistical model and handling the data. Because of the small size of the sample (and sample subgroups by schooling attainment), I excluded all who were severely handicapped or who had part-time jobs in 1965. Dummy variables for these characteristics are statistically highly significant in Rogers' work, but because the set of observations having these characteristics is very likely to be highly heteroscedastic, they are omitted here. I also excluded individuals for whom earnings data were unavailable or zero in 1950. These omissions reduced the total sample by 19 from 364. Rogers measured social class by a four-valued single variable, whereas I trichotomized the sample by

The IQ coefficients for high school graduates and college graduates with one degree are about the same for discounted earnings. However, the 1950 IQ coefficient for high school graduates is quite small, while it is relatively much larger for college graduates with one degree. The small IQ coefficient for high school graduates indicates that the opportunity cost of earnings while attending college differs little over a wide range of ability, while ability makes a larger relative difference promptly in the earnings career of college graduates. Clearly this provides some incentive for those with higher IQ to attend college than others.

The small sample and large standard errors leave one with some uncertainty about the precise magnitude of the effect of ability on earnings. This leads to the question: how prevalent are the gross positive associations found in Rogers' data between terminal schooling level, ability measures, and earnings in other samples? Tables 5 and 6 show means and standard deviations from two other samples that broadly confirm the pattern of Table 2.

The Project Talent sample statistics (Table 5) tend to confirm these broad associations. The log-earnings differentials are from two years early in the lifetime earnings profile. (Earnings in 1962 were obtained when most individuals in the sample were a year beyond high school graduation, except for the lowest schooling level.) The trivial inversion in earnings for E_1' and E_2' and the larger inversion for E_3' in 1966 are probably due to the relative labor force experience (and associated increase in productivity) for individuals from these schooling levels. If the sample is restricted to those reporting earnings in both 1962

dummy variables. Finally my results are based on coefficients from individual regressions by schooling level to allow for full interactions with the other variables. Rogers pooled his observations in a single regression in which the IQ-schooling interaction was allowed (but no other schooling level differences).

TABLE 5—MEANS AND STANDARD DEVIATIONS* OF LOG-EARNINGS AND ABILITY TESTS BY EDUCATION LEVEL (Project Talent)

Education Level	Sample Size <i>N</i>		<i>LE66</i>	<i>LE62</i>	<i>COO1</i> ^a	<i>COO4</i> ^b
	66	62				
<i>E</i> ₁ '	71	22	8.535 (.359)	7.813 (.449)	122 (54)	65 (27)
<i>E</i> ₂ '	698	293	8.527 (.413)	8.025 (.444)	159 (44)	89 (33)
<i>E</i> ₃ '	1232	532	8.313 (.556)	7.69 (.616)	191 (42)	119 (41)

* Numbers in parentheses are standard deviations.

^a A composite ability index that correlates well with IQ.

^b A quantitative ability measure.

Source: Large subsample of high school juniors taking the battery of Project Talent tests in the spring, 1960. Schooling levels are *E*₁' high school nongraduate, *E*₂' high school graduate, and *E*₃' two or less years of college. Earnings for 1966 are reported earnings for the preceding twelve months plus self-employment income (which was negligible for this sample). Earnings for 1962 were obtained by inflating a weekly or monthly pay rate to annual equivalent earnings. The 1962 pay rates were reported in a two-year follow-up questionnaire sent to members of this cohort. Earnings distributions are truncated by a procedure that eliminates extreme tails from the log-earnings distribution. The procedure essentially excludes observations over 2.75 standard deviations from mean log-earnings for the year. (The mean and standard deviations were computed from an initial calculation before earnings figures were excluded.)

and 1966, the (geometric) mean earnings of *E*₂' are about 12 percent greater than *E*₁' earnings in 1966. The combination of college attendance and military service leaves the *E*₃' cohort with relatively little civilian job experience even in 1966.

The Husén sample of a Swedish student cohort (Table 6) has a similar association between schooling, earnings, and ability with one important exception. At the time students in the Swedish sample attended school, the Swedish educational system was organized strongly in the continental tradition under which relatively few students obtain high levels of education. Most students terminated their formal academic schooling with the completion of the *folkskola* (elementary school) at the age of 14. The attrition rate of those starting the *realskola* (secondary school, which might be completed at the ages of 15 or 16) is high, and only a minority succeed in obtaining the secondary degree (or degree from more vocationally oriented alterna-

tives). Thus the IQ of those graduating at this level have the relatively high mean IQ of 111, and there is in this sample no tendency for further rise in mean IQ with additional schooling except for a small increase for *E*₁' (doctors and dentists). On the earnings side, it appears to take some years before those completing the *realskola* (or next level) to catch up with earnings of those with lower schooling attainment.

Turning next to the interaction between ability and earnings for a given level of schooling, only very preliminary calculations are available. However they seem to tell about the same story as the Rogers' calculations. In the observed sample range, measured ability has a negligible association with earnings for low levels of schooling attainment at all points along the life earnings profile. The squared simple correlations of both Project Talent ability measures with log-earnings for both years are less than .015 in all cases except for .09 for 1966 for the *E*₁' class.

TABLE 6—MEANS AND STANDARD DEVIATIONS* OF LOG-EARNINGS AND IQ BY EDUCATION LEVEL (Swedish, from Husén)

Education Level	Sample Size N^{**}	LE68	LE64	LE59	LE54	LE49	IQ 1938
E_1''	38	9.858 (.762)	9.621 (.459)	9.192 (.335)	8.905 (.313)	8.000 (.558)	80.5 (12.2)
E_2''	305	10.169 (.377)	9.766 (.335)	9.391 (.273)	9.017 (.194)	8.217 (.361)	95.6 (12.7)
E_3''	77	10.350 (.413)	9.941 (.307)	9.419 (.284)	9.038 (.247)	8.233 (.415)	102.5 (12.8)
E_4''	39	10.493 (.494)	10.032 (.292)	9.559 (.333)	9.034 (.194)	8.016 (.405)	111.0 (10.4)
E_5''	49	10.831 (.350)	10.170 (.419)	9.681 (.384)	9.101 (.218)	8.001 (.708)	110.5 (11.0)
E_6''	27	11.001 (.284)	10.308 ^a (.916)	9.946 ^a (.355)			110.4 (11.4)
E_7''	5	11.295 (.167)					114.2 (13.8)

* Numbers in parentheses are standard deviations.

** Sample size is for 1968 earnings. This number is also approximate sample size for 1938 IQ. Number of observations continually decline for earnings from earlier years.

^a Sample size is five.

Source: Earnings data obtained directly from individual tax returns. IQ data calculated from T. Husén data tape. Sample is based on all male Malmö third graders (mostly ten years old) who were tested for ability in 1938. Schooling levels are E_1'' less than *folkskola* (elementary school completion), E_2'' *folkskola* completion (usually at age 14), E_3'' some *realskola* (secondary school), E_4'' *realskola* completed (usually at age 16 or 17), E_5'' *gymnasium* (roughly junior college) completion (ages 19–21), E_6'' a university degree (excluding E_7''), and E_7'' doctor or dentist. Earnings distributions are truncated by a procedure that eliminates extreme tails from the log-earnings distribution. The procedure essentially excludes observations over 2.75 standard deviations from mean log-earnings for the year. (The mean and standard deviation were computed from a second iteration, using this rejection criterion.)

In the Swedish sample, initial calculations for the E_4'' level or less indicate that the 1938 measure of ability has negligible effect on earnings for the lowest three education classes. At the E_4'' (*realskola*) level, the 1938 ability test interacts more strongly with earnings (with squared simple correlations of .086 and .096 with log-earnings for 1959 and 1968). Further calculations are required to estimate ability parameters in equations similar to the Rogers' regressions for the higher education levels. The initial results suggest they will be very small for the lower levels of schooling.

REFERENCES

- O. Ashenfelter and J. D. Mooney, "Graduate Education, Ability, and Earnings," *Review of Economic Statistics*, February 1968, 50, 78–86.
- G. S. Becker, *Human Capital and the Personal Distribution of Income*, Woytinsky Lecture No. 1, University of Michigan, 1967.
- Z. Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," in W. L. Hansen, ed, *Education, Income, and Human Capital*, New York, NBER, 1970, 71–115.
- W. L. Hansen, B.A. Weisbrod, and W. J. Scanlon, "Schooling and Earnings of Low Achievers," *American Economic Review*, June 1970, 60, 409–18.
- T. Husén, *Talent, Opportunity, and Career*, Uppsala, 1969.
- D. C. Rogers, "Private Rates of Return to Education in the United States: A Case Study," *Yale Economic Essays*, spring 1969, 9, 88–134.
- T. W. Schultz, "Resources for Higher Education: An Economist's View," *Journal of Political Economy*, May 1968, 327–47.

DISCUSSION

WILLIAM T. GARNER: The papers by John Hause and Eric Hanushek are complementary: Hanushek deals with the production of school achievement, and Hause with the effect of achievement or ability as measured at some point on earnings thereafter. It will soon be time for studies which can unify these approaches and deal comprehensively with the development of human resources and economic behavior, whether viewed narrowly in terms of money-income streams or more broadly in terms of the allocation of time by individuals to various kinds of activities.

Hause's paper is itself a partial report on some research whose time has come. He is to be commended for seeking out and refining longitudinal, as opposed to cross-sectional, data for the study of ability and income. If my information is correct, Hause shares the credit for seeing to it that the Project TALENT survey follow-ups obtained information on the post-school earnings of students. That longitudinal survey might otherwise have been devoid of such economic information. Hause's research should form a valuable contribution to the understanding of the ability-income problem.

It is unfortunate that all four of the cohort samples he is working on could not be included in this report—even if the four groups are not readily comparable. What we are really waiting for in this area are some results, which in the case of the data reported today are admittedly weak. In addition, it would be good to have some idea of what is in the remaining data, so we can know what refinements to expect. Hause's treatment of "ability" in the present paper, for example, is too vague. He defines ability loosely as "learning power," but operationally he uses I.Q. scores. Such scores measure ability in at least two ways. First, they measure the achievement of certain kinds of skills, that is, the ability to perform certain tasks. Second, they measure the rate at which these certain tasks can be performed. Such a confounding inheres in most ability measures, not just in I.Q. scores. Furthermore, it is important to distinguish the ability to learn from other possible in-

terpretations of ability. Hause edges up to, but does not deal directly enough with analyzing whether employers prefer to hire abilities with respect to specific skills, or the ability to learn (if there is such an identifiable ability). If it will be possible to measure the effect of schooling on ability to learn, say if we can determine the pace of learning or changes in the rate at which value is added to individuals, we can improve our understanding of individual investment in school, in on-the-job training, and so on. It may be, for example, that schooling and ability may interact in favor of the more able. This would satisfy the apparent theoretical necessity (when returns to schooling are measured in monetary terms alone) that education give higher absolute marginal returns to the more able, in order for continued schooling to maintain a favorable rate of return despite the higher foregone earnings of the more able. The investigation of such possibilities will require at least one more observation per individual than seems to be available in the data reported today. That is, we would need at least two within-school "ability" measures, as well as post-school earnings.

Hause assumes that there are some "components" of schooling which do affect the earnings of individuals differentially exposed to them, but which of the several he mentions—cognitive achievement, socialization effects, etc.—which of these is most important he is not prepared to say. And he does not include learning to learn as a school component, as mentioned above. I would like to see more attention to this matter of school components, although I do not know if he has appropriate data, as a contribution to the study of school productivity. Econometric studies of this type, of course, are by their nature historical: in the Rogers data which Hause uses, for example, most men in the sample had received the bulk of their education by World War II. Even the most skeptical observer of the educational scene will surely agree that the process and content, not to mention the duration and institutional forms, of schooling have changed a good deal since then—for better,

worse, or naught. Insight into what particular aspects of schooling had been important in the past, however, would be helpful in the analysis of contemporary schooling and the many alternatives to formal schooling which appear to be taking shape.

Eric Hanushek is also to be commended for developing better data. We have already had more than enough attempts to estimate educational production functions with data collected for different purposes (such as Coleman Report or Project TALENT data) or, in the case of school records, sometimes for no apparent purpose at all. Hanushek has actually arranged for data to be collected with production questions in mind, and this alone is a step forward.

Hanushek's conceptual model may contain excess baggage, however. Consider his first statement, that the effects of schools and teachers cannot be looked at in isolation—that we must consider all of the factors that enter the educational process. This statement is used in educational production function studies, and this one is no exception, primarily to justify the inclusion of a vector of socioeconomic measures. Perhaps this is unnecessary. The socioeconomic measures are thought to represent primarily environmental influences on the child's ability to learn, to comprehend instructions, on his sharing of school or teacher values, on achievement acquired through home resources, etc. All of these influences, if they exist, are presumably embodied in the student and could be measured at that point. It is the student who interacts with school resources, not the socioeconomic abstractions. Is it not better, therefore, to measure student attitudes and abilities directly?

While Hanushek sets out to remedy some of the shortcomings of previous studies, he does not attack what are in my opinion the greatest defects of previous studies: the failure to measure the actual utilization of inputs, and to specify the technology in which that utilization takes place. Every production function study of schooling so far has in effect equated the presence of a variable with its utilization. But, as everyone here knows, it is the services of productive units, not the

units themselves, which are factors of production. Hanushek mentions an attempt to record the time of teachers in various activities, but little was done with this. Yet, strictly speaking, it is teacher behaviors, not teacher "characteristics," which are inputs to the schooling process. In addition, it is just as important to measure the time of students in various activities: reading a text, listening to teachers, working problems, or practicing whatever skills are to be developed. In general, the time of students with various configurations of school resources, human or otherwise, can be measured, and the students' production of defined skills can likewise be measured. It is necessary to realize that it is the student who produces the achievement. We speak of school production functions, but in an important sense it is the student who has the production function, which the school attempts to influence through compulsory requirements, subsidizing of certain inputs, and so on. A good deal of analysis based on individual achievement or human capital production functions has been carried out by Yoram Ben-Porath and Gary Becker. That work has been concerned primarily with the post-school investment behavior of adults, and has so far had little application to in-school learning.

Likewise, the relevant technology must be specified, so that we know not only how much of each resource was used, but how it was used. A problem of prior studies, from which Hanushek has not yet escaped, is that by aggregating over different curricula and technologies, few inferences about particular curricula or technologies can be drawn. The same is true in general economics, where Cobb-Douglas estimates of the relative factor proportions in the American economy give little useful information to the businessman who must decide how best to make his widgets.

The solution to these problems, I believe, lies in continuing in the direction taken by Hanushek toward micro data. But I think he has not gone far enough. There is little reason to stay within the naturalistic school setting, which has time and again been shown to abound with inefficiencies. Instead, I think the next step is to the experimental identification of educational production functions.

There are many examples in both industry and agriculture of the experimental identification of production functions, although we ought not to expect the education of children to be as amenable to study as the growing of corn or the manufacture of oil products. Experimental conditions allow, incidentally, for control of background (socioeconomic) variables by design rather than by regression techniques. Once we have some real knowledge about marginal productivities and marginal rates of substitution at hand, we can, I believe, contribute to the improvement of school practice.

ROGER ALCALY: "A child born today in the United Kingdom stands a ten times greater chance of being admitted to a mental hospital than to a university, and about one fifth of mental hospital admissions are diagnosed schizophrenic. This can be taken as an indication that we are driving our children mad more effectively than we are educating them. Perhaps it is our way of educating them that is driving them mad."—R. D. Laing, *The Politics of Experience*

The papers by Eric Hanushek and Herb Gintis provide an extremely interesting and instructive contrast, not only with respect to their analysis of educational production relationships but in their basic approach to the study of economic phenomena. Hanushek's paper represents the mainstream in both regards, taking the goals of the educational system as given and concerning itself primarily with the equity and efficiency of the system in meeting these goals. Gintis, on the other hand, addresses himself to certain fundamental questions about the economic role of education in society and the mechanism by which schools operate to fulfill this role. His findings cast considerable doubt on the economic relevance of traditional studies of so-called educational production functions, of which Hanushek's must be considered one of the more advanced.

Hanushek's paper reports an attempt to estimate "educational production functions" from a unique set of data, collected from a large public school system in California during the summer of 1969, which contains infor-

mation on educational "output" (reading achievement) and "inputs" at an individual level. These data represent a substantial improvement over those used in previous studies of this type (which relied primarily on the Coleman Report) and allow him to consider the "value-added" by the educational system rather than just the level of educational achievement, and to account for the cumulative effect of the school system, particularly individual teachers, in this process. In addition, his analysis is thorough and competent in a technical sense. Nevertheless, I feel very strongly, and Gintis' analysis supports my view, that this study and many others like it, are not addressed to the most important questions concerning the educational system in this country. Since many of these issues are recognized, at least implicitly, in Hanushek's paper but then dismissed, it seems clear that there are certain ideological biases inherent in the methodology employed and/or the reward structure of the profession and society which generate such research as the rule rather than as the exception.

Certain fundamental questions concerning the role and nature of the educational system (this is a broad generalization which does not take into account various differences which obviously exist between, for example, public and private schools) in society must be faced before a study which is concerned with the equity and efficiency of that educational system makes much sense. Are the schools fostering individual development in both cognitive and noncognitive dimensions or do they aim to produce well-behaved citizens and workers who will smoothly fit into the existing order of things regardless of the costs to themselves in terms of emotional development, personal satisfaction in life, and the ability to develop creative human relationships? And if the latter, as seems fairly clear from the increasing characterization of schools as jails which resort to drugs and other forms of inhuman treatment to keep students in line, are we happy with such an educational system and such a society? With respect to cognitive skills, do schools help students to understand the historical development and the basic nature of their society, or are students taught what amounts

to a rationalization of the existing institutions and power relationships? Again the latter seems much closer to the truth and we cannot avoid making a value judgment about the desirability of that kind of system in designing any research on education, whether the judgment be made explicitly or implicitly through an analysis of the equity and efficiency of the existing system.

Hanushek recognizes some of these issues early in his paper when he mentions cognitive development and socialization as the broad categories of educational "outputs" but dismisses consideration of the latter because of data availability. This, of course, is the typical tack taken in research of this kind by economists, but seems particularly disturbing in this case because the data were collected specially for the study. And although there is a great danger that the more easily quantifiable aspects of the educational system will be considered to the exclusion and confusion of the more basic questions referred to above, this need not be the case as Gintis' paper amply illustrates.

Gintis' analysis suggests, contrary to the commonly accepted view but consistent with the one outlined above, that "the economic productivity of schooling is due primarily to the inculcation of personality characteristics which may be generally agreed to be inhibiting of personal development," and that the "economic productivity" of schooling must be measured against an opportunity cost reflected in the repression of spontaneous, creative, and other 'anti-productive' characteristics in students." These findings are of critical importance and do considerable damage to much of the existing work on the economics of education, including "educational production functions," and welfare economics.

Even taken on its own terms—which as I have tried to indicate should not be done—there are certain basic shortcomings of the "educational production function" approach. Hanushek recognizes that all tests of hypotheses about the relationships between "inputs" and "outputs" are restricted to the limited ranges of values in the data sample. Consequently, the results can only be used to predict the effects of relatively limited changes in

the policy variables. And if, as implied in much of the previous discussion, the kinds of changes in the educational system that would be desirable are likely to be large rather than marginal, the policy guidance from these types of studies would probably not be too great. The problem is compounded even further by the fact that the estimated relationships which emerge from these types of studies are really average production relationships rather than true production functions since there can be no presumption that school systems were operating in an "optimal" fashion when the data were generated.

Hanushek closes his paper with the recognition that the results of his study (are they really results or merely hypotheses which must be tested with new data?) cannot be generalized since they are based on a small sample, and a suggestion that this prototype be repeated in other situations. Most of what I have said opposes this assessment. I think there would be a serious misallocation of resources involved in further studies of this kind. More work in the spirit of Gintis' analysis, however, is needed desperately.

CHARLES Z. WILSON: The papers presented at this session represent encouraging steps toward the use of sophisticated economic analysis in the field of education. Hanushek's paper in particular is exciting. Not only does he bring rigorous economic analysis to bear on a rather sticky problem—he produces some provocative findings as well.

Of special interest to me are the empirical foundations of the study. There is, in general, a reluctance on the part of economists to delve deeply into field work; the author deserves a measure of thanks for his effort. I am concerned, however, about his assumptions regarding the homogeneity of the school system. While it is useful to consider a school system as a micro unit, one must be extremely careful in making assumptions about uniformity in its behavior. I doubt whether potential biases from community specification variables are fully eliminated in microanalysis at the school system level. School systems, as a rule, serve many communities and there are times when these communities should be given explicit at-

tention in specifying educational processes. There are needs, in some instances, to recognize the usefulness of intraschool system data. Obviously this observation may be more relevant for large city school systems than smaller ones.

The author makes a major contribution to the development of educational production functions using historical or time-oriented data. Most researchers have used cross-sectional data. However, by using time-oriented data, the author raises the problem of providing measures of raw inputs or innate ability for his models. While I am not satisfied with the author's treatment of the problem, it is doubtful that a better alternative could be provided given the gaps in the notebooks of school psychologists.

The findings of the study should be accepted as a point of departure for further inquiry rather than as basic contributions to policy making. Educational administrators have felt for some time that the "pay parameters" of Hanushek's model may not be significant in the achievement of children and, hence, may be one of the driving forces behind the rising costs in our school systems. The educational return per dollar is embarrassingly poor. Thus, Hanushek's findings (that none of the attributes of teachers which are purchased have any effect on the quality of education) tend to confirm what many of us in the vineyards have suspected. Some would point to performance-cost ratios of higher education with the same concern. The "pay parameter" of a production function reflecting faculty input would not contribute materially to student learning. This is at the heart of our undergraduate revolution in higher education.

Returning to Hanushek's findings, his selection and analysis of teacher characteristics and their contributions to student achievement is far-reaching. The analysis is impressive, but one wonders if it is carried far enough. Why, for example, does he assume that Mexican-American data should be fitted to the same production model as the white sample? It would have been useful to test two versions of Equation (6): a white manual sample and a Mexican-American manual sample. I suspect, however, that the data was too

skimpy to support such an approach.

In general, the models of educational processes developed by Hanushek are simple and possess pleasing qualities for estimation purposes; they should not, therefore, be expected to be too descriptive. I am, nevertheless, concerned that the analysis may be too lean and could have benefitted from supplementary analysis of a noneconomic variety in selected instances.

Hanushek suggests we cannot be confident that any of the teacher attributes which are purchased by school systems have any effect on pupil achievement. This is questionable; we do not know enough about the characteristics of teachers vis-à-vis learning to make such a conclusion. A more reasonable statement would be that we cannot be confident about the effectiveness of the *traditional* teacher attributes. Considerably more work is needed on a microlevel to delineate a range of attributes that could be appropriately considered in a study such as Hanushek's. Like the multiple product firm, school districts frequently segment their markets. Customers are treated differently according to location, socio-economic status and their ability to process demands. In purchasing the services of teachers, school districts do buy color, sex, physical presence, and other qualities. Hanushek arbitrarily selected qualities of teachers to be considered explicitly. My question is: Why not determine, through a preliminary empirical inquiry of the district, school, or decision unit, the teacher characteristics to be explicitly recognized in the model?

After the tremors following Arthur Jensen's article, I am convinced that social scientists should be as exhaustive as possible in analyzing findings with perceived racial implications. Hanushek's findings—that teachers do not appear to count for Mexican-American pupils in the sense that changes in teachers and classroom inputs do not affect the achievement outcome of pupils—is incomplete and has perceived racial implications. As an economist, I share Hanushek's frustration over such findings; they place one's analytical and normative values in conflict. Such findings revive the fear that economics is powerless to handle complex human problems and, indeed, is a

"Dismal Science." As a discussant, I would like to exercise my prerogative as a casual observer and suggest one possible hypothesis for such findings.

I suggest that the "teacher don't count" hypothesis, with respect to the Mexican-Americans, stems from the feed-forward effects of negative cultural, teacher, and pupil attitudes. The initial achievement level variables A_{j1} or A_{j2} carry forward an accumulation of the composite effects of attitudes and other general learning experiences. Samuel Bowles hints at this type of problem in the following statement: "Student attitudes toward self and toward learning are consequences of past and present achievements (as well as other influences) and are important determinants of achievements." (*Conference on Research in Income and Wealth*, NBER, 1968, pp. 8-9.) The thought that negative attitudes or other socio-cultural factors may be present is revealed by Hanushek himself.

It is also interesting to note that no white collar or nonmanual component for Mexican-Americans appears in the study. Moreover, with classes containing as high as 63 percent Mexican-Americans, there was not a single Mexican-American teacher in the sample. These are conditions that might well contribute to poor self-image and to depressed motivation.

It should be noted that in addition to the fact the Mexican-American pupil may be attending schools that relate negatively toward his cultural values, the study measures achievement with the Stanford Achievement Test (SAT). This test is not very effective for Mexican-Americans who are culturally isolated in a school system. (This appears to be the case in the school system under study.)

Despite these criticisms, Hanushek has

done a remarkable job with a difficult and pioneering study. While one may not eagerly leap toward policy with his findings, the study opens some interesting areas for further inquiry. How useful is the value-added approach? Should we attempt to develop differentiated production functions for ethnic groups? What is the most critical set of teacher characteristics for student achievement: psychological, economic, or social?

Hause's paper contributes further to our appreciation of the value-added approach; it complements Hanushek's work. On the basis of preliminary findings, Hause suggests that ability or innate intelligence is not likely to be a separable factor in explaining lifetime earnings. In fact, ability may well interact with schooling and, hence, linear models are misspecified by not allowing for such interaction.

Gintis' paper suggests that economic and social status (variants of educational output) may be more related to personality and organizational than cognitive factors. In a sense, he goes beyond the input-output relations developed by Hanushek and Hause. This is helpful; affective variables have some desirable qualities that may be useful. For example, the explanation of Hanushek's findings with respect to the achievement of Mexican-Americans may be rooted in the interaction of affective variables. Beyond this general consideration, Gintis should be applauded for some rather sophisticated analysis. He goes somewhat too far in projecting the implications of his analysis.

While all three papers add to understanding, they point to the need for much more empirical and theoretical work. Thanks to the authors, Hanushek, Hause, and Gintis, researchers have a better sense of direction.

ROUND TABLE SESSION ON THE ACADEMIC LABOR MARKET

Whither the Market for Academic Economists?

By ALLAN M. CARTTER
New York University

Six years ago, as the result of several studies of the academic labor market, I made more enemies than friends by predicting that the faculty sellers' market, in which we had lived with only brief interruption since World War II, would end by 1969 or 1970. I foresaw an impending problem of over-supply of young PhD's seeking academic jobs in the 1970's and early 1980's, rather than the catastrophic shortage of faculty that was commonly predicted.

In 1969, after four years away from the subject, I returned for a reassessment, only to find that the market situation for the 1970's looked even worse than it had from 1964's vantage point. Few persons need persuading of this fact today after the last year of professional meetings and their somber evidence that good teaching jobs are indeed becoming scarcer. Federal cutbacks in research and development expenditures, overtaxed state educational budgets, and all the other factors Earl Cheit has identified as contributing to "the new depression in higher education" have exacerbated the situation. Some persons believe that the present imbalance in the academic labor market is a temporary phenomenon due to these latter factors. The projections for economics presented in this paper, however, indicate that we are likely to experience a continuing imbalance as far ahead as one can see due to

more basic demographic trends and institutional factors which make supply very insensitive to changing demand conditions.

Although the available data leave much to be desired, they provide adequate benchmarks to construct a picture of present trends. A broad survey of college faculties in 1964 by the Office of Education gives us a fix on faculty distribution in various disciplines. Table 1 shows the number of economists teaching in collegiate institutions, and the percent with the doctorate.

In Table 2 I have estimated the number of full-time economics faculty members for the period 1960-1980, using the 1963-64 base and assuming that the share of the enrollment market for economics remains constant over the period.

TABLE 1—FULL-TIME ECONOMISTS ON COLLEGE
AND UNIVERSITY FACULTIES, 1963/64*

	Full-Time Faculty		
	Total	with PhD	% PhD
Universities	2,031	1,485	73.2%
4-Year Colleges	1,559	796	53.1%
2-Year Colleges	270	23	8.5%
All Institutions	3,860	2,304	58.4%

* In addition there were 1,166 part-time faculty (about 400 f.t.e.), one-third of whom held the doctorate.

Source: Computed from James Rogers, *Staffing American Colleges and Universities*, Office of Education, 1967.

TABLE 2—ESTIMATED SIZE OF ECONOMICS TEACHING FACULTY, AND ADDITIONAL ECONOMISTS WITH DOCTORATE¹ REQUIRED TO MAINTAIN QUALITY OF STAFF

Academic Year	Full-time Teachers of Economics	Annual Increments to Teaching Staff	New Teachers with PhD Required		
			For Expansion	For Deaths & Retirements	Total Required
	(1)	(2)	(3)	(4)	(5)
1957/58	2,608				
1958/59	2,987	179	105	40	145
1959/60	3,118	131	77	41	128
1960/61	3,316	198	116	44	160
1961/62	3,476	160	93	45	138
1962/63	3,598	122	71	46	117
1963/64	3,860	262	153	47	200
1964/65	4,187	327	191	51	242
1965/66	4,661	474	277	55	332
1966/67	5,120	459	268	62	330
1967/68	5,531	411	240	68	308
1968/69	6,074	543	317	74	391
1969/70	6,360	286	167	81	248
1970/71	6,562	202	118	85	203
1971/72	6,831	269	157	88	245
1972/73	7,132	301	173	91	264
1973/74	7,374	242	141	95	236
1974/75	7,706	332	194	98	292
1975/76	8,045	339	198	103	301
1976/77	8,335	290	169	107	276
1977/78	8,593	258	151	111	262
1978/79	8,859	266	155	114	269
1979/80	9,089	230	134	118	252
1980-85 Aver.	9,525	82	48	125	173

This assumption approximates the experience of the last decade. The enrollment estimates and marginal student staff ratio used to project future faculty needs are drawn from my 1969 paper for the Joint Economic Committee.¹

Column 2 indicates the annual increments (two-year moving averages) in faculty needed to meet expanding enrollments. Column 3 indicates the number of new faculty with the doctorate required for expansion purposes each year to maintain the quality of faculty (i.e. 58.4% of column 2). Column 4 is an estimate of combined deaths and retirements of fac-

ulty holding the doctorate.² Column 5 of Table 2 estimates the annual need for economists with the PhD sufficient to maintain the present level of quality (as measured by highest degree). It is clear that requirements peak in the 1965-69 period, dip in the 1969-74 years, and steadily decline after 1975.

The basic reason for the decline in new faculty required is the decrease in the rate of growth of the 18-21 age group after 1968, only likely to be partially offset by rising college attendance rates. At the present time about 70% of high school graduates begin some type of formal post-secondary education, so we have nearly

¹See A. M. Cartter and R. L. Farrell, "Academic Labor Market Projections and the Draft," *The Economics and Financing of Higher Education in the United States*, Joint Economic Committee of the Congress, U. S. Gov't. Printing Office, 1969, pp. 357-74.

²The overall death and retirement rate, given the faculty age distribution, is about 1.81%. Since approximately 73% of faculty over age 50 are PhD's, I have used a 1.32% rate in column 4 (i.e., $.73 \times 1.81$).

TABLE 3—DOCTORATES REQUIRED TO MAINTAIN QUALITY OF FACULTIES,
AND NUMBER AVAILABLE FOR COLLEGE TEACHING, 1960-80

Year (September)	PhD's Granted	Number Available for Teaching*	Number Required*	Surplus	Surplus as % of Doctorates Awarded
	(1)	(2)	(3)	(4)	(5)
1958	239	159	145	14	6%
1959	221	147	128	19	9%
1960	237	158	160	(2)	(1%)
1961	266	177	138	39	15%
1962	268	178	117	61	23%
1963	331	220	200	20	6%
1964	385	257	242	15	4%
1965	410	273	332	(59)	(14%)
1966	458	305	330	(25)	(5%)
1967	546	364	308	56	10%
1968	600	400	391	9	2%
1969	656	437	248	189	29%
1970	717	478	203	275	38%
1971	763	509	245	264	35%
1972	861	574	264	310	36%
1973	916	611	236	375	41%
1974	985	657	292	365	37%
1975	1,055	703	301	402	38%
1976	1,118	745	276	469	42%
1977	1,188	792	262	530	44%
1978	1,280	853	269	584	45%
1979	1,352	901	252	649	48%
1980-85	1,500	1,000	173	817	54%
	Aver.				

* Available doctorates in any given academic year are compared with doctorates required in the next academic year. Most doctorates are granted in June, and these recipients become potential teachers in September of that year.

exhausted rising college attendance as a source of growth. We have lived so long focused on the population bulge that hit the colleges in the mid-1960's that we may forget that after 1978 the 18-21 age group will shrink by nearly 20% over the following decade.

Columns 1 and 2 of Table 3 suggest the supply situation. Between 1960 and 1970 the percentage of doctorates awarded in economics rose from 2.41% to 2.70% of all doctorates. I have assumed that economics will continue its relative growth to approximately 3% of all doctorates in 1980. The overall doctoral projections are from my 1969 paper, and are, if anything, rather conservative in comparison with

latest Office of Education and National Science Foundation projections.³

Traditionally about two-thirds of all persons receiving the doctorate in economics have entered academic employment. There is nothing sacred about this proportion, except that it suggests that attractive nonacademic employment opportunities have grown over the last fifteen years at about the same pace that

*My PhD projections for economics are almost exactly the same as Lindsey Harmon's most conservative projection in his companion paper. His supply figures are about 16% greater than mine for each year due to his inclusion of agricultural economics. I believe his projections of total PhD's for 1980 are over-optimistic in light of current trends—but he has a very good track record up to this point.

teaching positions have expanded. As a benchmark I have used 66.7% to apply to column 1 to estimate the available, or aspiring, supply of new college faculty with the doctorate as shown in column 2 of Table 3.

Columns 2 through 5 of Table 3, illustrated in Figure 1, compare the estimated annual number of economists with the doctorate available with the number required to maintain present quality. It is apparent that during the period of rapid expansion in the 1960's the market stayed surprisingly in balance. There were modest deficits in 1960, 1965 and 1966, but the picture dramatically and permanently changed beginning in 1969.

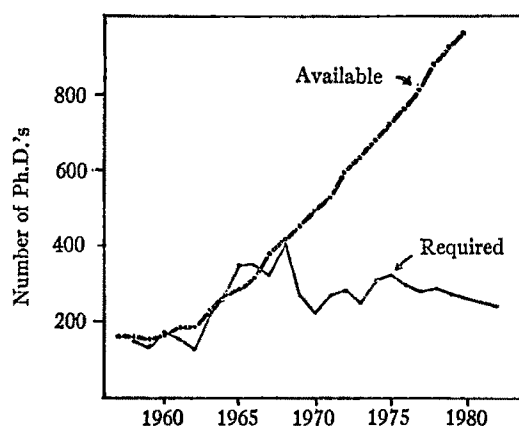


FIGURE 1

Economics Ph.D.s available for college teaching, and numbers required to maintain quality of faculty: actual and projected, 1957-1982.

Source: Table 3, columns 2 and 3.

Ordinarily the market adjusts reasonably well to short-run swings when the deficit or surplus is not more than about 15% of total doctorates available. When the imbalance is relatively small most of the available doctorates are hired by colleges and universities, and the percentage of total faculty with the doctorate improves or erodes slightly. Minor adjustments in salaries and recruiting zeal are probably

effective measures to clear the market. However, from 1970 onward, the anticipated surplus will be too large to be absorbed in the academic sector. Allowing for the fact that a number of new teachers will always be hired a year or two before they receive their doctorate, and that there will always be some distinguished persons who meet all criteria except the doctorate degree (e.g. the Kenneth Bouldings of this world), not all economists first hired will be drawn from the ranks of those with the doctorate. I suspect the maximum proportion, even with extremely high hiring standards, would be about 80% for universities, 75% for four-year colleges, and perhaps 25% for two-year colleges (where the teacher must be a generalist in social studies). With this optimal mix, and the distribution shown in Table I, 74% (instead of 58.4%) is about the highest proportion one might expect for new teachers with the doctorate, even in a strong buyers' market.

Next year's market is a good illustrative case. Table 3 indicates that about 245 economists with doctorate will be required just to maintain standards. If, instead, hiring standards were raised to the maximum, only about 65 additional Ph.D.'s would be required. But, as Table 3 indicates, there are expected to be more than 500 new doctorate holders seeking academic posts, leaving a surplus of nearly 200 even under the most ideal standards. This situation will now be a permanent phenomenon unless the structure of higher education changes or the number of doctorates awarded falls far below the anticipated numbers. In 1972-75 the available supply may drop 5-10% below the levels shown in Table 3 as the result of the draft lottery; these will be chiefly deferrals, however, and will not markedly affect either the short-run or long-run picture. The major impact of the draft so far has been to raise significantly the proportion

of women in graduate and professional schools—a desirable result in its own right, but not one likely to correct the aggregate imbalance of supply and demand.

A continuing over-supply of 30% or more will have a major impact upon academic labor markets, and in the long run is likely to have a depressant effect upon students' choices to go on to graduate school. For the next five years the projected doctorates are already in the pipeline, but beyond about 1975 the degree projections shown in Table 3 must be open to question quite apart from the uncertainties of the draft.

A few institutions, notably Harvard and Yale, announced decisions to reduce by about 25% their incoming graduate classes this year. At this stage, however, such reductions in private universities are more than compensated for by the rapidly emerging (predominantly public) universities that are just developing PhD programs. Given the nearly ten-year planning and implementation period, it is unlikely that expansion in public systems can or will be turned off in the coming decade. Thus a graduate education system that had the capacity in the mid-1960's to carry us to about 1980 will probably end up being 50–75% larger than the nation requires in the foreseeable future. In the case of economics this excess capacity may not be so costly, since it consists largely of staff (which can be diverted to undergraduate teaching duties) and library resources. In some of the scientific and technological fields, however, there is customarily a much greater fixed investment in facilities which may be underutilized.

It is improbable that the relative abundance of PhD's will result in unemployment; rather there is likely to be underemployment, and many who aspired to academic posts will have to take less attractive positions in government and industry,

or teaching posts in lower education for which they are ill-prepared. The more plentiful supply of PhD's can be a healthy factor for collegiate education, for departments can be more selective in employing and granting tenure to the above-average scholar-teacher. It may also provide some benefits for nonacademic sectors, for they will be able to employ many more persons with greater training and potential than they could have attracted in the past. It will result, however, in many personal disappointments in career paths.

For colleges and universities, the changed market conditions raise serious questions as to whether normal retirement ages should be rolled back to the early 60's with annual extensions for the most distinguished scholars. Unfortunately, with the annual surplus predicted, the retirement age would have to be about 50 to absorb the new entrants (i.e., the retirement rate would have to be about 5% annually, instead of the approximate 2½% rate that would be provided by current retirement policies with an even age distribution among faculty).

Six years ago I concluded my first excursion into this subject with the advice that 1970 was a good year by which to have achieved tenure. Reassessing the academic labor market for economists from today's vantage point, I can only offer the consolation to the young emerging economist that if he had invested his graduate years in the Victorian novel, 18th century Italian poets or the decline of the Ottoman Empire, he would have much greater cause for concern. At least in our profession one is not completely dependent upon continued academic prosperity for employment. And good economists will always be in short supply.

In a paper given at the AAAS meeting in Chicago. I urged consideration of a federal policy to identify and financially support 75 to 100 "national universities," and

to discourage the expansion or continuation of expensive doctoral programs in other institutions. Failure to have a positive federal policy, I firmly believe, could lead to a chaotic situation in the late 1970's and early 1980's.

The greatest danger that exists today is that federal agencies and state legislatures will react to excess capacity by merely reducing their support levels. Given the extreme financial pressures most institutions are already struggling under, and the inability of colleges and universities to reduce levels of output without soaring unit costs (chiefly due to faculty tenure), such a Malthusian kind of imposed adjustment process could be devastating. Earl Cheit's study due for publication in February under Carnegie Commission auspices, provides a clear foretaste of the difficult financial problems ahead.

I have made a strong recommendation to the President of our Association that a high level commission be appointed to review the state of the profession and to consider national policy guidelines with

respect to graduate education in economics. After spending a day at the AAAS meeting in Chicago, I regret to say that the various scientific professional societies are several years ahead of the economists—whose natural propensity it should be—in studying the market for their PhD's and the adequacy of their graduate programs for the 1970's.

The shoemaker's son traditionally may be the last to get a new pair of shoes, but it seems slightly absurd that economists should be the last to take a keen interest in their own labor market. I urge the Association to draw upon the varied talents of its membership to study appropriate steps that might be taken to minimize the painful readjustments that I believe will be required to overcome a near-permanent disequilibrium in the academic labor market. As an Association we shall have failed the current generation of students in economics if we ignore the problem and merely wait for a classical market adjustment to correct the projected oversupply of PhD's.

The Supply of Economists in the 1970's

By LINDSEY R. HARMON
National Research Council

I have been asked to present some data on projected output of economists over the next decade. I have also been asked to take a look at the projections made nine years ago at the AEA convention, which, as I understand it, is the last time previously that the AEA has taken a searching look at the economics manpower situation.

Sometimes in this business, grubby details trip one up seriously, so it is essential to spend just a moment on matters of definition, boring as that may be. In his 1961 projections, and in the data for the preceding period which he reviewed, Ewan Clague included both economics as reported by the U.S. Office of Education, and agricultural economics. The report in the AEA Journal didn't say so, and it is only by checking the numbers in the books on earned degrees conferred that one is able to discover this important fact. Actually, this is convenient for my report, because for most of the period over which we have data in the Doctorate Records File of the Office of Scientific Personnel, these two kinds of economists are combined. Beginning in 1969 we have a separate break-out for the agricultural economists. In our records, as distinct from the USOE practice, the person puts his own label on his field, and we accept it as long as it is within our code structure. If not, it goes into a miscellaneous "other" category, until that category grows large enough to require some new fields to be recognized. In our data, there are probably also some graduates of departments of business administration who consider themselves economists, whereas in the USOE figures, they are likely to be reported in one of the business administra-

tion categories, because the OE reports are compiled from registrars' data, which tend to follow departmental lines rather strictly.

With these preliminary understandings, let us take a look at what Clague foresaw in 1961 as likely developments over the next decade. He anticipated that economics (including ag economics) would comprise 1.3% of all Masters' degrees, and 3% of all PhD degrees. He used the USEO projections of total numbers of degrees at each of these levels as a basis for his calculations. It seems that this was his major mistake, but it is hard for one department of the government to tell another its projections are no good, even if they think so. I do not know what Clague thought in 1961, but the USOE record is one of chronic underprediction of output, particularly at the PhD level. We will come back to this matter of overall projections of PhD degrees a bit later. First for a look at Clague's projections set against the actual USOE output data for the years 1961-1968, which is as far as the data are available at this writing. Table 1 gives the essential information.

At the MA level, Clague had projected 1.3% of the degrees would be in economics; actually it turned out to be 1.4%. At the PhD level, he had assumed that 3% of the degrees would be in economics; actually it turned out to be 3.32%. If the projections of total degrees upon which he had based his calculations had been correct, he would have underestimated by about 7% at the MA level and about 11% at the PhD level, which I believe we would call pretty good shooting. His actual errors were about five times as great

TABLE 1—PROJECTIONS, 1961, BY EWAN CLAGUE, AND USOE ACTUAL DATA, OUTPUT OF ECONOMICS (INCL. AGRICULTURAL ECONOMICS) AT MA AND PhD LEVELS, ACADEMIC YEARS 1961–1968 INCLUSIVE

Academic Year Ending In	Clague Projections		Actual Degrees Conferred		Errors: Degrees Conferred Minus Projection			
	MA	PhD	MA	PhD	Raw Numbers		As % of Actual Degree	
					MA	PhD	MA	PhD
1961	1023	326	1129	373	106	47	– 9.39	–12.60
1962	1024	353	1217	379	193	26	–15.86	– 6.86
1963	1088	381	1369	433	281	52	–20.53	–12.01
1964	1179	394	1449	500	270	106	–18.63	–21.20
1965	1301	409	1597	538	296	129	–18.53	–23.98
1966	1359	434	1901	586	524	152	–28.51	–25.94
1967	1361	471	2147	680	786	209	–36.61	–30.74
1968	1426	512	2358	752	932	240	–39.53	–31.91
1969	1668	539	—	—	—	—	—	—
1970	1807	561	—	—	—	—	—	—
8-Year Sum. 61–68	9761	3280	13,167	4241	3406	961	–34.89	–22.66

at the MA level and about twice as great at the PhD level, and increased progressively over time, as the USOE projections of total degrees fell farther and farther behind actual output. So it appears that the major culprit, from the standpoint of foreseeing the future, is not the estimate of the share that economics will have of the pie, but the size of the pie. It may be interesting to remember this ten years hence, when someone else compares my forecasts with the actuality. I have some unorthodox approaches to the estimate of total PhD output.

When we take a look at the future, I will confine my attention principally to the PhD level, not because the Masters people are unimportant, but because I have no special knowledge of the MA output situation, and because I do not have any data since 1968 on which I might base projections. If one is interested in attempting projections from this point in time, my recommendation would be that the rate of increase experienced over the

last several years might give a reasonable clue as to the future. That growth rate has been about 8% per year. There was a discontinuity in the data about 1965 that appears to be a matter of definitions; before and after the discontinuity the growth rate was quite regular. Simply projecting the 1968 figure of 2,358 at 8% per year gives us 4,041 by 1975 and 5,938 by 1980. That's a guess, and about as good a guess as I would be able to make.

Let me reiterate the definitions I will use. In the data of the Office of Scientific Personnel, economics includes agricultural economics except for the year 1969, when there were 86 cases grouped in "agriculture, other" who were agricultural economists. When these are added to the 703 which our data show for 1969, we have a total of 789, which falls very close to the trend line for the rest of the 1960's. As I mentioned earlier, there are no doubt some graduates of business administration departments in our figures, which explains why they are a bit higher than the USOE

TABLE 2—PROJECTION OF ECONOMICS PhD'S BY TWO DIFFERENT PROCEDURES

Fiscal Year	N and %*	Set 1: Annual Percentage Increments				Set 2: Fixed Percentage of Total PhD Output					Projected Total by fiscal year (see text)
		7%	8%	9%	10%	2.6%	2.8%	3.0%	3.2%	3.4%	
		1	2	3	4	5	6	7	8	9	10
70	N %	844 2.96	852 2.99	860 3.02	868 3.05	741 2.6	798 2.8	855 3.0	912 3.2	969 3.4	28,500
71	N %	903 2.88	920 2.93	937 2.99	955 3.05	815 2.6	878 2.8	941 3.0	1003 3.2	1066 3.4	31,350
72	N %	967 2.79	994 2.87	1022 2.95	1050 3.03	902 2.6	971 2.8	1040 3.0	1110 3.2	1179 3.4	34,675
73	N %	1034 2.72	1073 2.82	1114 2.93	1155 3.04	988 2.6	1064 2.8	1140 3.0	1216 3.2	1292 3.4	38,000
74	N %	1107 2.71	1159 2.84	1214 2.97	1271 3.11	1062 2.6	1144 2.8	1226 3.0	1307 3.2	1389 3.4	40,850
75	N %	1184 2.68	1252 2.83	1323 2.99	1398 3.16	1149 2.6	1237 2.8	1325 3.0	1414 3.2	1502 3.4	44,175
76	N %	1267 2.64	1352 2.82	1442 3.01	1538 3.21	1247 2.6	1343 2.8	1439 3.0	1535 3.2	1631 3.4	47,975
77	N %	1356 2.62	1460 2.82	1572 3.04	1691 3.27	1346 2.6	1450 2.8	1553 3.0	1657 3.2	1760 3.4	51,775
78	N %	1451 2.57	1577 2.79	1714 3.03	1860 3.29	1470 2.6	1583 2.8	1696 3.0	1809 3.2	1922 3.4	56,525
79	N %	1552 2.51	1703 2.76	1868 3.03	2046 3.31	1606 2.6	1729 2.8	1853 3.0	1976 3.2	2100 3.4	61,750
80	N %	1661 2.45	1840 2.71	2036 3.00	2251 3.31	1766 2.6	1902 2.8	2038 3.0	2174 3.2	2309 3.4	67,925

* Percent Economics is of PhD total for the academic year.

data for economics plus agricultural economics. The average annual growth rate in the OSP data for economics, as above defined, is about 9.5% per annum. My guess would be that this is a reasonable growth rate to project for the future. Perhaps it is too high. In Table 2, I have given the results of several different projections by two general methods, each with a set of quantitative variations.

The first method, illustrated by columns 1 through 4, takes off from our most recent data, for fiscal year 1969, and pro-

jects a constant growth rate. This is a very satisfactory method over a few years where the growth rate has been reasonably steady and there are no clear reasons for anticipating a change in growth. The first four columns of the accompanying table project growth rates of 7%, 8%, 9% and 10% per annum from the common base of 789 economics PhD's in FY 1969. Below each figure in this table is a percentage entry, calculated as the percent economics PhD's are (in the projection) of all PhD's, based on a projected total of

PhD's indicated in the final column of the table at the far right.

This total PhD projection is the basis of the next five columns of figures, each of which assumes a different constant percentage of total PhD's in economics, ranging from 2.6% to 3.4% by .2% increments. Because the total PhD projection figures so largely in these data, it deserves special attention. The numbers you see here are a conversion to fiscal year data of a series I projected in June 1964, on the basis of data on the years 1920 through 1963. The original projection was made in calendar year terms because for 1920-1957 this was all we had. Conversion to fiscal year, which means July 1 through June 30 and corresponds rather well with the academic year, was made by multiplying the calendar year projection by .95, which was the average conversion value for the last several years. But how was the original calendar year projection made?

Of all the possible methods, I chose the simplest. That was to observe the actual performance of PhD output over a period that was not perturbed by war or its aftermath, or by economic depression. Such a period was found during most of the 1920's, a period characterized in certain histories as "the return to normalcy." For doctorate production this was true (however false it may have been for agriculture, for instance). During most of the

1920's, PhD output rose at about 10% per annum. Ergo, my base was 10%. But growth can't be expected to be steady; these curves have a way of fluctuating, and it seemed to me that from crest to crest I should expect about a 15-year interval. In 1963, the percentage increment was still rising from the 1957 low. So I guessed that it would rise a bit further, then begin to fall to a low of about 8% per annum in 1975, and rise again to 10% by 1980. This was simply a guess, based on past experience with *growth increments*, and involved the assumption of no major war and no severe depression. You will note that it did *not* involve an assumption about federal R&D dollars—the analogous period in the 1920's was not so fueled. So there you have the basis for Harmon's 1964 Series B. No revelation in a dream, no golden tablets, just a projection based on what seemed to me in 1964 to be reasonable expectations. So—how has the projection fared to date?

Actually, I have been exceedingly lucky. The original estimates were rough. The first year I rounded the calculation to the nearest 125, then for the next 3-4 years to the nearest 250, and thereafter to the nearest 500 cases. For the fiscal year conversion, I have simply entered the computer reading—i.e. the original rough guess multiplied by .95. The data are in Table 3.

TABLE 3—PROJECTED VS ACTUAL PhD PRODUCTION 1964-1970

Fiscal Year	Projection B	Actual Output (OSP)	Error	Percentage Error
1964	14,369	14,323	+ 46	+ .32
1965	16,150	16,341	-191	-1.17
1966	18,288	17,953	+335	+1.87
1967	20,425	20,385	+ 40	+ .20
1968	22,800	22,916	-116	-.51
1969	25,656	25,721	- 65	-.25
1970	28,500	29,436	-936	-3.28
Seven Year Total	146,188	147,075	-887	-.61

This track record is too good to be true, I know. I was simply lucky. But it's on the record, and the experience with this projection gives me no reason to substitute any other, or to change the 1964 guesses for any more recent ones. It still seems likely to me that we will have neither a major war (mobilizing millions) nor a severe depression during the 1970's. Quite obviously, atomic bombs would render these projections not only obsolete but irrelevant. As to the depression—well, you're economists and I'll not carry coals to Newcastle.

Given these sets of figures, which projection of economics PhD's seems most reasonable? Which one gives the best basis for planning? You will note that in Table 2 there is a general progression from column 1, the most conservative projection, to column 9, the most optimistic (if you want high output). Yet 3.4% of PhD's in economics is a high proportion, based on the past decade's actual experience. From 1958 through 1969, the percentage has varied from a high of 3.53 in

1961 to a low of 2.73 in 1969. By three year intervals, the averages over these twelve years were as follows: 3.25, 3.30, 3.14, 2.87. This shows a generally decreasing proportion over the past eight years, and perhaps a further decrease is "blowing in the wind." If so, the problem of how to place the PhD's would seem to be minimized. But also, if so, it means that economics as a discipline will be declining in its quantitative significance in the total academic picture. Not necessarily in power and influence—medicine has not become less important because the medical cartel has restricted output for a couple of decades. But then the realities of a cartel situation need to be examined. I have some comments on that subject, but my assignment was output projections, rather than policy, and adjustment of inputs in 1970 to possible markets in 1977 or 1978 gets me over onto other people's turf on this panel, so perhaps we should listen to them and I'll hold my fire for now.

Recent Behavior of Economists' Salaries

By FRANCIS M. BODDY
University of Minnesota

In this brief paper I will be using some of the information compiled in a series of annual surveys of departments of economics (now in its fifteenth year). The fifty-four departments surveyed (as of 1970-71) include all those (36) with published reputational ratings in the 1965 Cartter report, and practically all departments producing PhD's at a rate of five or more per year in the most recent years for which information is available.

A parallel survey of some selected departments of agricultural economics begun in 1962 with nineteen departments, (twenty-four are being surveyed in 1970-71), gives additional and comparable information in this field.

Since my assigned topic is "Recent Behavior of Economic Salaries," I must warn you that I will be using these most recent surveys to give an up-to-date picture of a large, but specialized and nonrepresentative sector of the market for economists. These survey economics departments granted the bulk of the PhD's in economics in the United States (more than 85 percent of the total in recent years). Similarly the departments of agricultural economics surveyed granted three-quarters or more of the PhD's in that field in recent years. These departments in a very real way set the tone for the new PhD's view of the academic and nonacademic market place as he faces it on the completion of his studies.

The fresh PhD, largely living in the student environment of these major departments, does not, in most cases, move into this same type of departmental environment, as a teaching and/or research scholar, as he moves to the faculty side.

Of the some 1,100 PhD's in economics who received their degrees in 1966-69 from the responding institutions (of a total of some 1,800 granted in the U.S.) and whose institution of placement was reported, eighty-five percent received their degrees from Cartter-rated departments, but only twenty-four percent were placed in such departments, and forty-three percent went to unrated U.S. academic departments. For the two highest rating groups combined, these sixteen departments granted forty-eight percent of the 1,100 PhD's, but took on to their own group staffs only fourteen percent of the total placed. The view from the departments in the survey is then a very special view, and one much more representative of the supplier/seller's view than the demander/buyer's view.

The pattern of placement of the supply from the survey departments of economics in 1966-1969 can be generally summarized as follows:

1. Almost 80 percent are placed (or return to) academic institutions (including a little more than 10 percent going to Canadian or other foreign institutions).
2. The remaining (22 percent of the total) nonacademic placement is dominated by the U.S. Government and the Federal Reserve System (38 percent); foreign banking, industry and government (20 percent); and U.S. research and consulting firms (16 percent). U.S. banking and industry, and international organizations (about 10 percent each) make up most of the remaining nonacademic placements.
3. Academic placements show a strong "downstream" pattern. Each "rated" group of departments tends to place their PhD's largely in departments of

TABLE 1—SALARIES OF ECONOMISTS—U.S.—1964–1970

	1970	1969	1968	1967	1966	1965	1964
N.S.F. Survey and Groupings:							
All PhD's Median			15,800		13,500		12,100
All Economists Empl. in Educ. Inst.			12,400*		10,500*		10,100
Economist-Primary Work:							
Teaching			12,000*		10,000*		9,700
Basic Research			14,000		12,200		11,000
All Economists Top Decile			24,000		23,000		20,000
Departmental Survey: (Academic Salaries for Year* beginning in 1900):							
Median Prof.	22,195 ^p	20,719	19,324	18,537	17,413	16,214	15,204
Median Ass. Prof.	15,626 ^p	14,992	14,175	13,359	12,449	11,823	10,993
Median Ast. Prof.	12,762 ^p	12,259	11,563	10,842	9,956	9,421	8,876

* 9–10 month academic year.

^p preliminary estimates, based on returns of 42 (out of 54) departments.

lower reputational rating, with some placement in departments of the same class or “upstream,” but with only very modest “upstream” placement from the “good,” “adequate plus” or “other” (nonrated) departments.

Turning now to the salaries of economists, I would like first to look to some comparisons with the salaries (earnings) of noneconomists both as to comparative levels and time trends. Table 1 shows some selected data on economists' salaries from the three NSF surveys, and from the departmental surveys. Other studies—particularly the study published by this association in 1968—have shown the consistency with which economists' salaries lead (or nearly lead) those of other academic-scientific fields in earlier years.

The departmental survey salary averages may be compared with the average rank-salaries of the general faculties within similar institutions, such as the “100 leading doctoral institutions” covered by the AAUP national annual surveys. These comparisons come out with quite consistent results. Economists' current salary levels are about 10 percent

above the general institutional level, and this represents a substantial increase over the more modest salary advantages shown by comparisons of the salary levels four or five years ago.

With the BLS annual survey of white-collar salary data, and the average weekly earnings in total private nonagricultural employment, some interesting comparisons of growth and trends in these data and in the academic economists' salaries are possible. In the recent five-year period (1964–65 to 1970–71) the pattern can be summarized as:

1. Over the five years, and particularly in the first three or four years, economists' salaries rose more rapidly than those of the other groups, typically by about two percentage points more per year.
2. Preliminary data from the survey shows a real slowdown in academic-economists' salaries (for 1970–71) ranging down to two percentage points below the growth rates of 1969–70, while the most recent white-collar survey shows a continuing general *rise* in the rates of increase through 1970.

The average weekly earnings in private

nonagricultural employment increased by twenty-six percent from 1964-65 to 1969-70, while academic salaries in the survey institutions increased from 33 percent (professors) to 40 percent (fresh PhD's) over this same period. It might be noted that the rise in the consumers' price index was 21 percent, much of this rise in the last few years, and its current rate of increase (about six percent) is now well ahead of the current rise in academic-economists' salaries.

The recent trends in academic-economists' salaries are summarized in Table 2, and (in growth rates) in Table 3, for the periods 1964-65 or 1959-60 to 1969-70, with some preliminary estimates for 1970-71. The recent growth rates are quite similar for the economists and the agricultural economists, and give little ev-

idence of anything but relative prosperity, although the corresponding rates of growth in the CPI show how real income growth has been slowing down recently. The preliminary returns for 1970-71, however, suggest that the current and prospective immediate future trends are less cheerful. These preliminary figures may be particularly disturbing to the new entrants onto the present market. Not only do they show a smaller rise in the hiring salaries for fresh PhD's in 1970-71, But also the departments' *opinions* as to what they expect to have to pay for 1971-72 entrants show little evidence of any continued rise. Also disturbingly, for these entrants at least, in recent years these forecasts have been remarkably accurate when measured against the salary levels of new appointments in the following

TABLE 2—SALARIES OF ACADEMIC ECONOMISTS 1964-65 TO 1970-71
(9-10 month academic year)

	1970-71	1969-70	1968-69	1967-68	1966-67	1965-66	1964-65
<i>In Major U.S. Academic Departments of Economics</i>							
Professor:							
Top Decile*	—	27,100	25,300	24,150	23,000	21,700	20,200
"Superior" (Top ½)	26,670 ^p	25,001	23,394	21,780	20,466	19,222	17,803
Median	22,195 ^p	20,719	19,324	18,537	17,413	16,214	15,204
Median Assoc. Prof.	15,626 ^p	14,992	14,175	13,359	12,449	11,823	10,993
Median Asst. Prof.	12,762 ^p	12,259	11,563	10,842	9,956	9,421	8,876
Median Fresh PhD's	12,167 ^p	11,842	11,012	10,486	9,780	9,250	8,675
<i>In Selected Major U.S. Departments of Agricultural Economics</i>							
Professor:							
Top Decile*	—	20,750	19,500	18,900	17,800	16,700	17,000
"Superior" (Top ½)	20,627 ^{p*}	19,332	17,848	16,972	16,384	15,244	15,176
Median	17,328 ^{p*}	16,681	15,513	15,120	14,425	13,125	12,528
Median Assoc. Prof.	14,507 ^{p*}	13,725	12,738	12,044	11,248	10,796	10,202
Median Asst. Prof.	12,509 ^{p*}	12,032	11,135	10,196	9,610	8,882	8,804
Median Fresh PhD's	12,247 ^{p*}	11,888	11,551	10,578	9,519	9,125	8,497

^p Preliminary estimates, from Table 1.

* Top Decile of all individuals in the reporting departments in this rank.

^{p*} Preliminary estimates, based on 16 (of 24) departments of agricultural economics reporting.

TABLE 3—AVERAGE ANNUAL RATES OF INCREASE
IN SALARIES OF ACADEMIC ECONOMISTS

	1969-70 to 1970-71	1968-69 to 1969-70	1964-65 to 1969-70	1959-60 to 1969-70
Economics Departments				
Superior Professors (Top $\frac{1}{3}$)	3.3 ^p	6.9	7.0	6.0
Median Professors	5.7 ^p	7.2	6.4	5.7
Median Associate Professors	4.2 ^p	5.8	6.4	5.6
Median Assistant Professors	4.9 ^p	6.0	6.7	6.1
Fresh PhD's, Median	3.2 ^p	7.5	6.4	6.5
				1961-62 to 1969-70
Agricultural Economics Departments				
Superior Professors	6.1 ^{p*}	8.3	5.0	6.0
Median Professors	5.5 ^{p*}	7.5	5.9	6.0
Median Associate Professors	7.6 ^{p*}	7.7	6.1	6.3
Median Assistant Professors	5.1 ^{p*}	8.1	6.5	7.0
Fresh PhD's, Median	5.7 ^{p*}	2.9	6.9	7.0
Annual Change in Consumers Price Index				
		5.8	3.8	2.6

^p Preliminary estimates based on 42 (of 54) departments reporting; comparison is of the 1970-71 and 1969-70 data of these departments only.

^{p*} Preliminary estimates from Table 2, rates compare only the departments reporting in both years.

year. The listener can perhaps decide whether or not this is an example of a self-fulfilling prophecy.

From the survey data, some of which is summarized in Table 2, some patterns of salary structure in the larger and well established departments emerge. A comparison of the averages of departmental averages for the "superior" (top $\frac{1}{3}$) and "average" (lower $\frac{2}{3}$) of the faculty in each rank (which is roughly a range of 50 percent of the persons, between the lower $\frac{1}{3}$ and the upper $\frac{1}{3}$ in each rank in each department) shows:

1. In the professor rank the "superior" salary level is 25 to 30 percent above the level of the "average" group in departments of economics, and from 20-25 percent above in departments of agricultural economics, and these ratios have been relatively stable over at least the last 8-10 years.

2. In the associate and assistant professor ranks the similar ratios of "superior" to "average" group salary levels run generally from 9 to 14 percent over the last five years, and show no real trends over the full 8-10 years of the data.

By tabulating the data by "Cartter rating" groups, a rather clear pattern develops:

1. Most strongly for "superior professors," but also for "median professors," there is a steady and substantial increase in salary levels as one moves from nonrated up through the rating categories, with a differential of 30 to 40 percent and more between the extreme groups in recent years.
2. For the associate professor rank no clear pattern emerges. In the most recent years there are little evident differences among the groups; in earlier years the higher-rated department sal-

ary levels were only slightly higher than those of the other groups.

3. For assistant professors and "Fresh PhD's" there are but modest differences among the groups with little pattern to the differences, except that the highest rated group has distinctly and consistently *lower* salary levels for both the assistant professor and the fresh PhD's than do the other departmental groups in the surveyed departments.

Some further items from the recent survey that may give some useful input to short-run forecasts of the market, and of academic salary levels for new entrants, are these:

1. After practically no change in the output of economics PhD's on the part of these survey institutions from 1967-68 to 1968-69, preliminary and partial data shows an increase from 1968-69 to 1969-70 of about 25 per cent in this most recent year. The agricultural economics departments' output, however, has not increased.
2. Scholarship / fellowship support reached a peak in the survey departments in 1967-68 but has declined only modestly in the last two years (down about 6 percent for 1970-71).
3. Teaching / research assistantships peaked in 1968-69 and have fallen very modestly in 1969-70 and 1970-71 (down 4.2 percent in 70-71).
4. Graduate enrollment in economics has held practically level for the last four years in the survey departments and showed practically no change for 1970-71 over 1969-70.

The sudden burst of output of PhD's in 1969-70 by the major economics departments may therefore be the result of a combination of episodic factors: the first effects of a conscious effort, backed with expanded support programming, to regularize and hasten the completion of the degrees of students in process a few years ago; the softening in the market for "not quite completed" PhD's that most departments reported last year; and the resulting feeling on the part of both the depart-

ments and their advanced students that the students had better complete their programs and get on the market before it softens still further. So far there seems, however, to be but few signs of the feared massive reductions in student support programs, and enrollments in the pipeline have shown no current decline in the major producing departments.

A slowdown in growth of salaries of academic economists and the leveling of modest decline in enrollments and in support programs for graduate students in economics, combined with the slowdown in the general growth of academic institutions which provide such a large part of the traditional and current demand for economists, may be the warning signals of a period in which the accelerator principle may suggest a substantial reversal of the patterns of the recent decade or more of fast growth in both supply and price for our products. One danger is, of course, that departments, students, and institutions may overreact to the current situation and to the prospects for the immediate future, to the neglect of the longer-run view of the needs and opportunities in the field. Another is that the number of new and untested departments that have been expanding doctoral programs in economics will have their products coming onto the market during a period when competition for placement will be even more intense, and the resulting feedback may result in a less than optimum adjustment of supply capacities in and among institutions.

The market is changing, and the first clear signals are now coming through above the noise-level, but what they signify for the immediate future is not yet clear as to the magnitude of the changes, or the length of the new era they suggest is coming. Nor is it clear what institutional or departmental responses either will be or should be in this year of changing trends. Nineteen seventy-one will be an interesting year!

Improving the Operation of the Academic Labor Market

By DAVID G. BROWN
Miami University

The charge issued to me by our panel chairman was to be prescriptive, to accept the statement and analysis of market problems in other presentations, to concentrate upon specific suggested solutions. Before turning to suggested solutions, a brief description of the market is appropriate and necessary.

Description of the Market for Academic Economists

Economists' services are contracted in a patchwork market broken by obscure intermediaries where partly irrational choices are made from partial information. Market perfection is inhibited by factors such as ethical dates of resignation, antipirating pacts among universities, folklore condemning extensive mobility, promotion from within, inbreeding and outbreeding, tenure, nontransferable fringe benefits, the difficulty of evaluating a position before occupying it and the difficulty of evaluating a candidate separate from the environment in which he will be working, the geographic separation of homogenous job opportunities, the time and money costs in job search and change of residence, the highly personal value scales of both buyers and sellers, the absence of a recognized intermediary, the reluctant maiden myth in job search, the subjectivity of job choice, and the superspecialized character of demand.

To portray the market for academic economists in hard data, let us revert to the five-year-old data summarized in Tables 1 through 3.

Without attempting to analyze why, let

me point up some of the more obvious observations that may be drawn from these tables.

Compared with other professors, job-changing economists are slightly younger (lines 5 and 6). Economists have a and less likely to hold a PhD degree (lines 5 and 6). Economists have a greater tendency toward research (lines 17 and 25).

Most mobility is initiated by economists, not their employers (line 18). Turnover is relatively high (line 26).

Economists tend not to be joiners and convention goers (lines 8 and 9).

Compared to other professors, economists are well paid by income (line 27), academic rank (line 23), and teaching load (line 24).

In economics more than in any other discipline, salary and future salary prospects are extremely important in job choice. When choosing between their two best job offers, economists tend to ask equally "what will I be doing in my next job?", "what resources will be available to help me do my job?", and "how much will I be earning?" (lines 19 and 20).

Although many channels are pursued in searching for a new job (line 12), informal contacts from buddy-to-buddy and mentor-to-intern usually yield the best result (line 13). The reluctant-maiden tradition still dominates academia, although many candidates have had amazing success with "cold turkey letters," convention placement services, and commercial teachers agencies.

Job changing usually means residence changing, with 50 percent of all professors moving more than 500 miles (line 14). The job search and move consumes about three weeks of productive time. Negotiations are firmed up in early April. Candidates typically pay moving expenses. The economists are slightly

TABLE 1—PERSONAL CHARACTERISTICS OF NEWLY HIRED FACULTY,
4-YEAR COLLEGES AND UNIVERSITIES, 1964-65

Line no.	Characteristic	Economists	All Social Scientists	All Non-Scientists
1	Age (median in years)	33	34	35
2	Age at Leaving Graduate School	30.2	30.5	n.a.
3	Percent Male	94%	91%	78%
4 ^a	% Protestant	n.a.	43%	56%
	% Catholic	n.a.	12%	15%
	% Jewish	n.a.	8%	6%
	% No Religious Preference	n.a.	31%	19%
5	% Ph.D. Holder	44%	50%	40%
6	% Ph.D. Holder (student previous yr.)	26%	33%	n.a.
7	% No Publications	60%	51%	62%
8	Prof. Assn. Memberships (mean #)	1.8	2.3	2.5
9	Conventions attended per year	0.8	1.0	1.1
10 ^a	% student in previous year	48%	43%	39%
	% faculty in previous year	35%	40%	36%
	% teacher below college prev. yr.	1%	2%	5%
	% Bus. or govt. in prev. yr.	9%	5%	5%
11 ^b	% loosely committed to teaching	9%	11%	n.a.

Source: David G. Brown and Jay L. Tontz, "The Mobility of Academic Scientists," A Report to the National Science Foundation, 1966.

n.a. means not available.

^a Vertical totals do not equal 100% due to omission of "other" answers.

^b A man is loosely committed to teaching if he did nothing to look for a college teaching job, was not a college teacher in 1963-64, and would have accepted a job outside college teaching if his current job were not available.

more successful in negotiating employer-paid moves and in closing contracts at an early date.

In spite of the many market intermediaries, more than one-third of all job changers rate the market mechanism as "poor" or "very poor" (line 21).

These data are for 1964-65. Immediately ahead we can foresee more forced and less free mobility, wider choice for employers and fewer options for job hunters, more nonprofessional constraints upon hiring, and more unionization. If the much discussed oversupply fails to materialize because the economists who are surplus to traditional markets are ab-

sorbed by new employers (for example, economic analysts in industry, social science teachers in high schools) and because surplus economists create their own demand (for example, founding and selling consulting services that previously weren't demanded), the market balance and behavior is likely to remain about the same. If, on the other hand, the oversupply does materialize we can expect a shift of employer effort from searching for candidates to selecting among those who appear on the doorstep, and a shift of candidate effort from selecting among several acceptable alternatives to finding at least

TABLE 2—CHARACTERISTICS OF JOB CHANGE, NEW FACULTY,
4-YEAR COLLEGES AND UNIVERSITIES, 1964-65

Line no.	Characteristic	Economists	All Social Scientists	All Non-Scientists
12	Searched new job through—			
	% former professor	n.a.	52%	40%
	% classmate/personal friend	n.a.	26%	26%
	% college placement office	n.a.	28%	43%
	% departmental office (grad sch.)	n.a.	44%	35%
	% convention placement service	n.a.	22%	16%
	% cold turkey letters to dept. chr.	n.a.	43%	47%
	% journal want ads	n.a.	15%	5%
	% commercial teachers' agency	n.a.	5%	9%
	% did nothing	22%	33%	33%
13 ^a	Found best job offer through—			
	% former professor	17%	18%	13%
	% classmate/personal friend	17%	16%	17%
	% college placement office	5%	6%	8%
	% departmental office (grad sch.)	14%	11%	6%
	% convention placement service	7%	4%	2%
	% cold turkey letters to dept. chr.	17%	14%	20%
	% journal want ads	1%	3%	1%
	% commercial teachers' agency	2%	2%	3%
	% did nothing	18%	23%	28%
14	% moved more than 500 miles	n.a.	51%	45%
15 ^b	Productive days lost in job change	n.a.	13 days	10 days
16	% moves where college paid expense	22%	18%	13%
17	Median date of job acceptance	Apr. 6	Apr. 12	Apr. 24
18 ^c	Could not have remained at previous job (in percent)	21%	21%	17%
19	% rejecting teaching job with higher pay	n.a.	16%	11%
20 ^d	Weights assigned when deciding between two best job offers—			
	% courses to be taught	28%	44%	35%
	% research facilities	8%	11%	9%
	% teaching load	7%	12%	11%
	% competency of colleagues	8%	9%	11%
	% salary	19%	7%	8%
	% future salary prospects	8%	4%	6%
	% stature of school	4%	5%	6%
	% quality of students	5%	4%	4%
21	% Evaluating job finding possibilities as "poor" or "very poor" instead of "good" or "excellent"	n.a.	38%	36%

Source: Same as Table 1.

^a Vertical totals do not equal 100% due to omission of "other answers." Other methods of job search include denominational placement, services, professional assns., the public employment service. Less than 5% search jobs by any one of these methods, and less than 1% find their best job through these channels.

^b Includes time looking for and at jobs as well as actual moving time.

^c The base is the number of new faculty who were also college teachers in 1963-64.

^d Factors not listed account for roughly 15% of the weights, though no one factor which is not listed accounts for more than 2% of the weights. The unlisted factors are congeniality of colleagues, administration and administrators, academic rank, fringe benefits, opportunities for outside income, nearness to graduate school, nearness to friends and relatives, climate, and cultural opportunities.

AMERICAN ECONOMIC ASSOCIATION

TABLE 3—CHARACTERISTICS OF NEW JOBS, NEW FACULTY,
4-YEAR COLLEGES AND UNIVERSITIES, 1964-65

Line no.	Characteristic	Economists	All Social Sciences	All Non-Scientists
22	% hired as associate or full profs.	17%	17%	16%
23	% of emerging PhDs hired above the instructor rank	98%	86%	n.a.
24	Teaching load (median in hours)	11.2	11.3	12.0
25	% spending more time in research than teaching	14%	14%	8%
26 ^a	Reason for vacancy—			
	% expansion of department	43%	49%	41%
	% predecessor went to another college	23%	25%	23%
	% predecessor left college teaching ^b	15%	13%	18%
	% predecessor want on leave	8%	4%	7%
	% Unknown	12%	8%	11%
27 ^c	Index of Annual income			
	All new faculty with PhD	151	144	142
	All new faculty without PhD	119	114	108
	Emerging students with PhD	131	123	120
	Emerging students without PhD	113	108	100
28	% expecting to remain more than 3 years	36%	37%	46%
29	% teaching in schools with—			
	Less than 1,000 students	15%	14%	15%
	More than 5,000 students	34%	39%	40%
30 ^d	Region of job—			
	% North Atlantic	n.a.	30%	28%
	% Great Lakes & Plains	n.a.	30%	31%
	% Southeast	n.a.	18%	19%
	% West and Southwest	n.a.	22%	22%

Source: Same as Table 1.

n.a. means not available.

^a Vertical totals do not equal 100% due to omission of "other" answers.^b Includes death, retirement, and movement to nonacademic employment.^c Includes salary, consulting, and income from part-time jobs. Excludes dividends, interest, gifts, royalties, and sale of manuscripts.^d North Atlantic: conn, me, mass, nh, ri, ver, del, DC, md, nj, ny, penn. Great Lakes: ill, ind, mich, o, wis, ia, kan, minn, mo, neb, nd, sd. Southeast: ala, ark, fla, ga, ken, la, miss, nc, sc, tenn, va, w va, West: all others.

one job. In an oversupply market job switching is likely to be less casual and more at the initiative of the employee.

Suggestions for Improving the Market

The suggestions for improvement that follow proceed from the premise that more information is good, that a market mechanism that causes buyers and sellers

to be more fully informed, and to think that they are more fully informed, is to be preferred. In other words, it is assumed that societal welfare is enhanced by achievement of objectives by individual buyers and sellers and that buyers and sellers will come closer to achievement when better informed.

In the employment process there are

TABLE 4.—ALTERNATIVE SUGGESTIONS FOR IMPROVING THE MARKET FOR ACADEMIC ECONOMISTS

<i>Suggestion</i>	
I.	TO HELP COLLEGES LOCATE AVAILABLE CANDIDATES—
Ia.	Candidate should send Cold Turkey letters.
Ib.	Graduate mentors should send unsolicited letters of nomination.
Ic.	Graduate departments should prepare lists of available candidates to carry to conventions.
Id.	Slave market dossiers should be made available by mail.
Ie.	An academic registry of all economists should <i>not</i> be initiated.
II.	TO HELP CANDIDATES LOCATE AVAILABLE POSITIONS—
IIa.	AEA should publish free all academic vacancies.
III.	TO HELP COLLEGES EVALUATE CANDIDATES—
IIIa.	Candidate should list 15 professional acquaintances.
IIIb.	Candidate should have maintained a full placement file.
IIIc.	Hiring department should make at least 2 phone calls.
IIId.	When on campus candidate should ask to interview faculty from other departments, junior faculty and students.
IIIe.	Tenure should never be granted to a new hiree.
IIIf.	Department should assess its highest priority need (teacher, researcher, community contact, etc.)
IIIg.	Department should request copies of publications, evaluations of teaching, etc.
IIIh.	Decision to make offer should involve full department.
IIIi.	Departments should consider the perspective and candor of evaluators.
IV.	TO HELP CANDIDATES EVALUATE POSITIONS—
IVa.	Candidate should request roster of department members, with an indication of date of hire, publications, and training.
IVb.	Candidate should consult guidebooks, university literature, including town information, local newspaper.
IVc.	Candidate should phone a friend who knows the department, but is not a member of it.
IVd.	<i>AER</i> should run annually a two-page feature on conditions in the market, including salaries, fringes, institutions censured by AAUP, etc.

two stages: first, locating alternatives and, second, evaluating the alternatives located. At both stages there is both a buyers' perspective and a sellers' perspective. Thus, there are four points at which to improve the market—locating candidates by buyers (category 1), locating vacancies by candidates (category 2), ranking candidates by buyers (category 3), and ranking appointments by candidates (category 4). The alternatives within each of these categories as well as my recommendation are summarized in Table 4.

With one dollar plus two hours, a candidate can inform his "most favored" potential employers through a cold-turkey letter of his availability. This direct approach is efficient and proven effective.

About 50% of all candidates write at least one such letter, 20% find their best offers through cold-turkey techniques. An improved variation on the cold-turkey theme is to ask graduate mentors to write unsolicited letters of nomination to a candidate's ten most favored potential employers. When possible, a candidate should arrange for these letters to be sent between professional friends.

My major recommendation for improving the market mechanism is the publication of a full list of academic vacancies available as a regular feature of the *American Economic Review*. Publication should be free. Listings should be actively solicited, especially from the most prestigious graduate schools. No vacancy

should be published more than once. Institutions should be named (no anonymous boxes). Only positions in higher education should be included. Positions in junior colleges should be listed separately. A signed commitment to list departmental vacancies from the chairman should be a prerequisite for publishing an article by a faculty member in the department in the *AER*.

Providing this service would be good not only for the market and morale but also for the Association. As jobs become more scarce, economists will become more concerned about how jobs are filled. Each man will want his "fair chance" at those openings that do occur. Pressure will be—indeed already is—upon the Association to accommodate to this need. The publication of vacancies is not expensive, is equitable, and places the main burden of contact on the individual (the very man who has the greatest personal stake).

You will note that I am not advocating the use of a registry in this market that is turning toward favoring the employer, because the registry mechanism places too great a burden on the employer (the very party that has the least to gain in the excess-supply market) and is too expensive relative to the benefit. Thus the advocacy for the listing of academic vacancies is a strong one. If this step is taken, candidates will be sufficiently informed of job openings. No other step is necessary.

The current scene is a bad one. Each graduate school has a placement office that is receiving announcements of job vacancies. In each school these vacancies are coded, classified, and finally communicated to the ten or fifteen students who might be interested. The duplication of effort throughout the country is shameful, inefficient, and ineffective. A centralized list is needed, a list that is fully inclusive

and available to all. Such a listing will enable placement offices in graduate schools to concentrate upon counseling instead of administrivia. Moreover and more importantly, a centralized listing will be more inclusive, nondiscriminatory, and most informative. Such a list will be valuable not only to candidates who are seeking information about specific job alternatives but also to others who are seeking general market information concerning salary levels and the state of the market.

When informed of a vacancy, the candidate can then initiate his evaluation procedures. Here it is wise to gain a general impression of an institution by consulting one of the many guide books available to entering college freshmen (for example, Cass and Birnbaum). Letters to the public relations office and the chamber of commerce will return more specific information. Salary ratings such as those published in the Summer Issue of the *AAUP Bulletin* provide still another dimension. If these preliminary investigations are positive, the candidate can make contact and when on campus examine job subtleties and the adequacy of support facilities such as the library, computer time, laboratory equipment, and travel funds. It is wise for candidates to develop a check list for comparing competitive jobs.

In summary, to reduce the imperfection in the operation of the labor market for academic economists, I am recommending a consolidated listing of vacancies available. This new market instrument, coupled with the numerous suggestions regarding the collection of additional information by both candidates and employers, should improve the quality of decisions made in this market and result in a better allocation of the nation's resource that is the services of economists.

A Younger Economist's Views on the Market*

By ROBERT P. STRAUSS
University of North Carolina

The recent turnabout to a buyer's market particularly affects new PhD's and I shall concentrate on the two issues I think are most important to us: the "crisis" in the market and its roots in the expectations of new graduates, and the reward system in the market which new academics subsequently face.

The "Crisis" and Its Expectational Roots

The rapid increase in the number of new PhD economists and the recent slowdown in the growth rate of academic vacancies has forced increasing numbers of new PhD's to take first jobs in government and industry. The majority of first jobs are still in educational institutions, but this percentage has recently fallen.¹ (See Table 1.) Also, it seems likely that new PhD's are being forced to lower prestige institutions than previously. Nevertheless, while type of employer has changed, as Professor Boddy indicates in his paper, very few PhD economists are actually unemployed.

The much-discussed "crisis" arises, then, not from a failure to obtain employ-

ment but rather from disappointment about first jobs and unsatisfied expectations about the start of one's career path. These expectations in turn are determined by a common set of values most of us hold, values which are implicitly or explicitly acknowledged when we chat about what it takes to be a "successful" economist. I think it is useful to articulate these values, for they not only explain current disappointments but also explain why new PhD's may continue to be disappointed in the future.

Professional bliss for most economists seems to be employment at a high-status university and a research career, only lightly sprinkled with teaching responsibilities. Evidence of a successful career is a lengthy bibliography, appointment to A.E.A. committees and the *A.E.R.* editorial board, and now even the chance of a Nobel prize. Put more directly, a "good" economist is a research academician at a big-name university who stresses publication and identifies with other research economists.

What accounts for this widespread emphasis on academic affiliation? Given that an individual wants to pursue a research career, it seems reasonable to argue that opportunities for research can be found outside the university. Admittedly, the intellectual stimulation of an academic department may be missing, but it would seem on balance that government and industry do have challenging research opportunities and the advantage of seeing one's analysis put to use.

* This paper benefited materially from the comments of Richard Schramm, Cornell University.

¹ There is a good deal of variation among disciplines in the percentage of first jobs that are in educational institutions. Generally, the humanities send the highest percentage back into academe; 86.7 percent of the 1969 doctorate recipients in philosophy went to educational institutions. In the social sciences the figure was highest for sociology (79.6 percent) and lowest for psychology (50.2 percent), while in the sciences the majority of new PhD's did not go into academe. Respective figures for engineering and chemistry were 28.3 percent and 17.7 percent.

TABLE 1—DISTRIBUTION OF FIRST JOB BY TYPE OF EMPLOYER
OF NEW PHD ECONOMISTS: 1967-69
(Percent)

Fiscal Year	Educational	Government	Industry	Non-profit	Unknown	Post-Doctoral Fellowship	Total Number	% Female
1967	67.8	7.4	4.4	3.0	15.4	1.7	700	5.4
1968	68.3	6.4	5.0	3.1	15.4	1.3	756	4.6
1969	64.8	9.3	7.1	3.7	12.3	2.4	799	6.1

Source: National Research Council, *Summary Report* (1967, 1968, 1969): Doctorate Recipients from U.S. Universities (Washington, D.C.: Office of Scientific and Technical Personnel).

Resistance on the part of new PhD's to nonacademic first jobs seems to come from two sources: fear that the trip out of academe is a one-way journey and a suspicion that gaining an audience for one's research in general economics journals (such as the *A.E.R.*) requires an academic affiliation. Since academicians produce PhD's (presumably in their own image), they are perhaps naturally reticent to hire someone outside the academy, someone who by initial or subsequent job choice has not supported the value that the academy is the place to be. The self-fulfilling nature of this value tends in turn to insulate the university from the rest of the world and perhaps gives rise to student charges of "irrelevance." A second implication of this value is that by keeping capable young people in academe, a technological and institutional gulf is created between those who make policy and those with theoretical and methodological expertise.

The suspicion about lack of publishing opportunities for nonacademics and especially government economists has been corroborated by Stigler² and more recently by Coe and Weinstock.³ Stigler

found for a time-series sample that academicians contributed about 95 percent of the articles in general economics journals compared with about 3 percent for government economists. Coe and Weinstock found for a more recent cross-section that academicians authored more than 90 percent of the papers written in economics journals. This disproportionate representation of academicians in the journals which they control is not surprising, for given the above set of values, almost by definition a nonacademic is out of the professional mainstream.

An indirect effect of the emphasis on research is that universities, presumably in the business of education as well as the production of knowledge, hire individuals who spurn teaching. New faculty members frequently attempt to minimize their teaching load and presumably the time spent in preparation per course. Both of these activities militate against the university effectively educating its students. Of course, the research emphasis benefits the university to the extent the scholar attracts grant money which pays for overhead and salaries. However, when grants become scarce and tuition income relatively more important, departments find themselves overstaffed with people who do

² See George J. Stigler, *Essays in the History of Economics*. (Chicago: University of Chicago Press, 1965), p. 45.

³ Robert K. Coe and Irwin Weinstock, "Editorial Policies of Major Economics Journals," *Quarterly*

Review of Economics and Business, Vol. 7 (Winter 1967).

not want to teach and who no longer pay their own way.

If we look at the way PhD economists are trained, we find that they not only are imbued with an antiteaching ethic, but, even worse, are not prepared for teaching. Graduate students who do any teaching typically teach freshmen and painfully learn by doing, not by direction.

While the ethic in the profession may be research and publication, in point of fact a majority of academic, PhD economists surveyed in 1965 had not published at all. Of the 3,800 economists with PhD's at educational institutions and who are in the 1966 *National Register of Scientific and Technical Personnel*, 61 percent had nothing cited in the *Index of Economic Journals*. A second observation regarding the research ethic is that while graduate students may be told about the desirability of research and perhaps shown how to do it, many still fail to complete the dissertation, presumed evidence of research competence.

There are several explanations for both of these contradictions between professionally-espoused values and actual behavior. Few economists may wish to do research, few may be capable of it, and/or few may be able to "get along" with the right people to succeed in completing a dissertation or getting a paper published. What does seem clear is that while the value may be to do research, there are many economists who do not actually respond to it.

The Markets for Jobs and Ideas in Academe

The last section discussed the impact of the standard professional values on new PhD's and the extent to which many have actually responded to them. This section examines two kinds of imperfections in the academic labor market which particularly affect younger economists: entry re-

quirements and restrictions on the airing of ideas.

There are several entry requirements affecting the market for academicians. First, there is the obvious, occupational license, the PhD, that is now required of university faculty. Failure to have a sufficient percentage of PhD holders can lead to loss of accreditation which in turn keeps graduate schools in business. Other important entry requirements which can be identified include: race and sex screening, political screening, and profession value screening.

The clear absence of black academic economists is due to several factors. First, there are relatively fewer black BA's than white BA's for each age cohort, so we would expect there to be relatively fewer black graduate students and, hence, black PhD economists. This disparity in percentage going to college in turn is due to unequal provision of primary and secondary schooling, financing problems that reflect not only the high incidence of black poverty but also discriminatory treatment in capital markets, and discriminatory admissions policies in many colleges and universities. Also, previous discrimination in the labor market for black college graduates has created well-known income differentials⁴ which tend to discourage college-going. Secondly, those blacks who do get BA's and aspire for more training typically go into law or medicine which are not only more lucrative than academe, but also less prone to arbitrary administrative behavior. It would seem then that this pronounced absence of black academic economists can be explained by various types of discrimination prior to the academic labor market.

⁴Most recent evidence on this is in Ritchie H. Reed and Herman P. Miller, "Some Determinants of the Variation in Earnings for College Men," *Journal of Human Resources*, Vol. V (Spring 1970), pp. 177-190.

While women may or may not be discriminated against upon entry to graduate school, they do find academicians reluctant to hire them upon graduation. My female colleagues explain that for those with a PhD today the principal screening device is marital status. Apparently, departments are now (but were not previously) willing to hire single women on the same basis as men. Departments discriminate against married women because they are more apt to move for nonprofessional reasons; i.e., their husband's job may change and thus force the department to incur greater search costs. This screening, which may or may not be discriminatory, depending upon how one views this search-cost argument, may still militate against women if equity considerations weigh when salary and rank decisions are made. Since married men have greater income needs owing to greater family size, they will receive greater equity pay and/or experience a faster rate of promotion than the single women who are hired.⁵

The second type of entry restriction relates to political views and activities. Legislative reaction to campus unrest has forced some departments in state universities to consider seriously the political inclinations of young and prospective faculty since budgetary reprisal is a very real threat. I think the pay-pause for all college and university faculty in the California system speaks for itself.⁶ We have all heard rumors of similar pressures in the Midwest, East and South. Reprisals not only affect university budgets but also certain aspects of academic freedom. I think

the pressures are such that those candidates who look potentially embarrassing will simply not be considered for employment.

Coupled with this second entry requirement is the likely stricture that prospective employees have broadly conventional professional opinions. By this I mean one holds the professional values discussed above and economic views within the intellectual boundaries of Chicago and Harvard. For those who find a neo or post-Marxian ("radical") orientation more persuasive than conventional economics, the welcome in academe is apt to be subdued if not hostile.⁷ I recommend the Bronfenbrenner-Davis⁸ exchange in the *A.E.R.* a decade ago to those who have forgotten the McCarthy pressures on leftist faculty and suggest that the current wave of "conservatism" that is sweeping the country will have a chilling effect on nonconventional economic speculations in colleges and universities.

Having survived the above entry requirements, the young economist must find an audience for his research efforts. The second set of imperfections in the market place involves the publishing mechanism per se. Younger scholars, under pressure of tenure considerations, may be more paranoid on this matter than older scholars. However, there is evidence which substantiates the common complaint that journal editors publish works

⁵ Some current research I am doing with W. Lee Hansen and Burton A. Weisbrod indicates that in 1965, academic, female PhD economists earned \$3,000 less in salary and \$5,000 less in income than similar men. This result was obtained after correcting for research output, years since PhD and prestige of academic employer.

⁶ See *Newsweek's* (November 23, 1970) article on the California system.

⁷ Of note is Bronfenbrenner's recent article (Martin Bronfenbrenner, "Radical Economics in America: 1970," *Journal of Economic Literature*, VIII, September, 1970) which encourages the profession at large to hire radical economists. One wonders if Carnegie-Mellon will promote the ultimate irony by hiring those who view Messrs. Carnegie and Mellon as villains rather than heroes.

⁸ Martin Bronfenbrenner, "Notes on Marxian Economics in the U.S.," *American Economic Review*, LIV (December, 1964); Horace B. Davis, "Notes on Marxian Economics in the U.S.: Comment," *ibid.*, LV (September, 1965); and Martin Bronfenbrenner, "Notes on Marxian Economics in the U.S.: Reply," *ibid.*

TABLE 2—FOUR UNIVERSITIES' (%) SHARE OF TOTAL PAGES IN THREE ECONOMICS JOURNALS (*A.E.R.*, *J.P.E.*, *Q.J.E.*): 1950-69

	<i>A.E.R.</i>		<i>J.P.E.</i>		<i>Q.J.E.</i>	
	1950-59 ^a	1960-69 ^a	1950-59 ^b	1960-69 ^a	1950-59 ^b	1960-69 ^a
Total Pages Published by Journal	5663	7207	4652	7471	6291	6486
University (%) Share of Total Pages						
Chicago	3.8	3.8	16.6	10.6	—*	3.2
Harvard	3.8	3.6	—*	1.9	14.5	12.3

^a From Cleary and Edwards, *op. cit.*, Table 3, p. 1013.

^b From Yotopolous, *op. cit.*, Tables 1 and 2, p. 667.

^c From Siegfried, *op. cit.*, Tables 1, 2, and 3, pp. 14-16.

* Less than 100 pages.

of their departmental colleagues and students first, works of their other friends second, and works of others last. The usual reply to this charge is that those departments house the most productive scholars and their excellence explains their success in publishing. However, since productivity is essentially a necessary condition for excellence, there is no direct way to ascertain if this argument is more than circular. Two kinds of evidence can be provided against it and in favor of a favoritism model of the publishing mechanism—evidence on the structure of the mechanism which creates the possibility that favoritism may occur and direct evidence on who publishes where.

Since most journals (including the *A.E.R.*) retain the author's name when sending a manuscript out for review, there is at least the possibility that matters other than merit may weigh during the review process.

Direct evidence on favoritism comes from several sources. Crane⁹ compared institutional affiliations of authors in the *A.E.R.*, which retains the author's identity for review, and in the *American Sociological Review*, which removes author's

identity prior to review, and found that the institutional representation in the *A.S.R.* was more diverse than in the *A.E.R.* A second set of evidence comes from Cleary and Edwards,¹⁰ Yotopolous¹¹ and Siegfried.¹² They tabulated for various years the institutional affiliation of authors in the *A.E.R.*, *Journal of Political Economy* and the *Quarterly Journal of Economics* and found that the name schools dominated these three journals. When we look more closely (Table 2) at who publishes in the *JPE* and *QJE*, we find that Chicago and Harvard dominated their own journals when compared to space each "won" in the open market (*A.E.R.*).¹³ In sum, there appears to be persuasive evidence which supports the charge of favoritism.

The impediments to successful partici-

⁹ Frank R. Cleary and Daniel J. Edwards, "The Origins of the Contributors to the *A.E.R.* During the Fifties," *American Economic Review*, LI (March, 1961), No. 1.

¹⁰ Pan A. Yotopolous, "Institutional Affiliation of the Contributors to Three Professional Journals," *American Economic Review*, L (December, 1960).

¹¹ John S. Siegfried, "Institutional Affiliation of Authors of Economics Papers: 1960-69" (unpublished manuscript, University of Wisconsin, August, 1970).

¹² The nonrandom pattern of university shares in Table 2 could be explained in a sense by a nonrandom pattern of submittals. However, if one believes authors send their papers to journals which they expect will publish them, then the submittal pattern may merely reflect the underlying pattern of favoritism.

⁹ Diane Crane, "The Gatekeepers of Science: Some Factors Affecting the Selection of Articles for Scientific Journals," *The American Sociologist*, II (November, 1967).

pation in the academic market place are several and perhaps no different in kind from those in other labor markets. Sex and race discrimination are everyday facts of life in our larger society, and getting ahead in the world is often based on matters other than productivity or merit. This observation, that academe reflects general societal practices, is not intended to comfort. Next I offer, by way of conclusion, a series of recommendations to improve the climate in the academy.

Conclusions and Prescriptions

When we combine overall market forecasts with the above themes, we find a scenario that is not very favorable to young economists. Certainly within academe the rate of promotion and institutional mobility of new PhD's will decline as the excess supply to academe grows during the 1970's. I anticipate a further tightening up of the research ethic vis-à-vis tenure at a time when teaching loads are likely to rise because of the grant squeeze. Editors are apt to face larger tides of manuscripts¹⁴ and young faculty wives are apt to see less of their husbands. Also, I expect the clampdown on political activism on campus to continue.

The "trickle down" of young economists will benefit the less prestigious institutions and government and industry as they hire more capable individuals than previously. However, to the extent that young PhD's are more politically active and to the extent that they enlarge the technology gap between younger and older faculty, there are likely to be costs as well as benefits.

What sort of actions can the profession take to make the market for economists more efficient and equitable and simulta-

neously promote the goals of science? In terms of the initial entry restrictions to academe, it would seem that department chairmen should begin obeying the laws with regard to sex and race discrimination. It should be noted that failure to mend one's way is no longer costless—Michigan has been threatened with loss of federal support for alleged sex discrimination in hiring.

For those economists who go to government or industry for employment, there will be problems of expectational adjustment as noted earlier. It seems to me that the avenues to and from academe should be broadened and incentives created to augment the flows. To this end, universities might do well to allow rather long leaves of absence to allow scholars to get involved with a real-world research problem and simultaneously hire industry and government economists to teach about economics in action. Along these lines, the A.E.A. can certainly better organize labor market information for permanent and temporary positions and attend to the related information problems of the future level and composition of demand for PhD economists. With regard to the demand issue *per se*, I think we will find that more applied economists are demanded for non-academic positions, and I would hope that the current panic about employment will not lead graduate institutions to cut back supply until it equals the vacancy rate in academe. The trickle down will have socially beneficial effects.

In terms of the market for ideas, I think there are several very clear things the A.E.A. can do. Certainly it can guarantee anonymity in the review process in its own journal and bring pressure to bear on the other economics journals published in the U.S. to follow suit. Secondly, the Association should increase the number of sponsored journals and encourage more specialization in each. I can envision, for

¹⁴ Coe and Weinstock, *op. cit.*, suggest that the average article acceptance rate of domestic economics journals has fallen, 1966 vs. 1957.

example, a theoretical journal, an empirically-oriented journal, and an economic policy journal, each published quarterly and sent to the membership.

It is patently ridiculous for an association numbering 18,000 members to publish only one association journal.¹⁵ Sociology publishes five and psychology even more.

I have not offered any prescriptions on the matters of political and professional

screening. It is in these areas that I think reform is most urgent but most difficult. The pressures on the academy to toe the middle of the political road are real and growing; the solution is to convince the funders of higher education of the efficacy of political diversity. With regard to the professional screening issue, I would argue that economics as a profession must diversify if it is to grow out of its current rut. And by this I mean not only must we tolerate in our journals rather different sorts of economics, but the intellectual health of the profession requires that such diversity be actively encouraged.

¹⁵ The Association of course publishes more than one journal; however, only the *American Economic Review* regularly considers unsolicited manuscripts for publication.

URBAN GROWTH AND DEVELOPMENT

Jurisdictional Fragmentation and Residential Choice

By BRYAN ELLICKSON
University of California

The United States, once a nation of small and relatively isolated towns, is now dominated by large metropolitan areas. When cities were isolated, most households lived in the political community in which they worked, but the growth of metropolitan areas and improvements in transportation technology have loosened this bond. In most metropolitan areas, households find a wide variety of political jurisdictions within feasible commuting range. Preferences for local public services can now be expressed through choice of residence as well as through the ballot.

Economists, for the most part, have analyzed residential location and local government as if they operated in complete isolation from one another. Most theories of residential location fail to recognize even the existence of local political jurisdictions. Theories of fiscal federalism—those theories which argue that optimal city size is determined by the trade-off between economies of scale and the diseconomies of decreased consensus—abstract completely from the question of why households decide to live within a particular jurisdiction.

Although we have no satisfactory theory of urban local government, economists have not been reluctant to propose reforms in existing institutions. Advocates of metropolitan government suggest that decisions must be made at the metropolitan level if externalities are to be internalized

and economies of scale realized. In contrast, the proponents of decentralization argue that further political fragmentation is required in order to provide greater variety in local public services. The only consensus, if any exists at all, is that present institutions of local government are inefficient. However, nowhere in the literature do we find an explanation of why, in view of this inefficiency, change is so rare. Annexations to the central city, relatively common at the turn of the century, ceased rather abruptly in most metropolitan areas after 1918. Subdivision of the larger political jurisdictions in our metropolitan areas does not appear at all likely. Thus, it seems reasonable to ask of an adequate theory of metropolitan political economy an answer to the question: why are existing jurisdictional boundaries so impervious to change?

To answer this question we must investigate the impact of local governmental structure not only on allocative efficiency but on the extent of redistribution from rich to poor as well. In a recent study of suburbs in the Philadelphia metropolitan area, Williams, *et al.* [5] report that, when heavy expenditures were involved, wealthy communities were unwilling to enter into cooperative agreements with less wealthy communities. Only when their wealth was about the same would cities agree to engage in a jointly financed program. Across the United States, proposals for metro-

politan government have typically met strong resistance from suburbs, particularly the wealthy suburbs which fear a diminution of tax base and a lowering of the quality of local public services. Development of a model of metropolitan political economy capable of explaining these aspects of urban political behavior is the principal aim of this paper.

The Nature of Local Government in an Urban Environment

The distinctive element that sets an urban political unit apart from the isolated frontier community is the mobility of its constituency. As Tiebout [4] observed more than a decade ago, households in an urban environment can "vote with their feet." In the urban community, unlike the isolated town, we cannot assume that the constituency is determined by forces exogenous to the political process. The location decision will be influenced by the public services and tax rates available in alternative jurisdictions. But once households move into a jurisdiction, they comprise the constituency which determines the tax and expenditure program that will be adopted. Thus, the nature of local government in an urban environment is the product of the simultaneous interaction of residential location decisions and the local political process.

This simultaneity raises serious problems for the empirical study of residential location patterns or local governmental decision-making in a metropolitan context. Oates [2] has recently provided some empirical support for the "Tiebout hypothesis," but—while he recognizes that governmental expenditures and taxes must be treated as endogenous—he treats median household income in the community as exogenous. Once the full implications of simultaneity are realized, the identification problems inherent in the application of regression analysis to the processes of

residential location or local political decision-making (at least in cross-section) become all too evident. Good empirical analysis must build on good theory. However, with the exception of the pioneering efforts of Rothenberg [3], there is virtually no theory of the simultaneous interaction of residential location and local political decision-making.

A General Equilibrium Analysis of Metropolitan Political Economy

I have suggested that recognizing the ability of households to choose among several jurisdictions is an essential prerequisite to understanding metropolitan political economy. Of course, households are not perfectly mobile; accessibility to employment imposes a significant constraint on the choice of residence. However, in order to explore the implications of urbanization for the nature of local government, I will adopt the extreme assumption that the location of households is not constrained by accessibility to jobs.

Household Preferences

In particular, assume that a household evaluates each jurisdiction solely in terms of the amount of housing, public services and other commodities available to it in that community. The maximum utility which the i th household can achieve if it resides in the j th community will be a function of the price of housing (p_d^j), the quality of public services (g^j), the tax rate (τ^j), the price of all other goods (p_c) and the household's wealth (w^i):¹

$$(1) \quad \phi^i(p_c, \bar{p}_d^j, g^j; w^i)$$

where $\bar{p}_d^j \equiv p_d^j(1 + \tau^j)$ represents the gross

¹ The function $\phi^i(\cdot)$ is the indirect utility function derived by maximizing the direct utility function, $U^i(c_i^j, d_i^j, g^j)$, subject to the budget constraint, $p_c c_i^j + \bar{p}_d^j(1 + \tau^j) d_i^j = w_i$, where c_i^j and d_i^j are the amounts of the composite commodity and of land consumed by the i th household if it resides in the j th jurisdiction.

price of housing, assuming that taxes are raised through a tax proportional to the value of housing. The price of the composite commodity, p_c , is assumed invariant throughout the metropolitan area. Land is assumed to be the only input into the production of housing; the land is owned by agents living outside of the metropolitan area, and it is supplied perfectly inelastically within each jurisdiction. I assume that a household chooses to reside in the community where it can attain the highest level of utility: i.e., the jurisdiction j for which $\max_j \phi^i(p_c, \tilde{p}_a^j, g^j; w^i)$ is attained.

Given the price of the composite commodity and the wealth of the i th household, the combinations of gross land price and level of public service yielding a particular level of utility, \bar{u} , must satisfy the following equation:

$$(2) \quad \phi^i(p_c, \tilde{p}_a^j, g^j; w^i) - \bar{u} = 0$$

Equation (2) defines an implicit relationship between \tilde{p}_a^j and g^j which I will refer to as a *bid price curve*. Bid price curves for households with wealth w_1 and w_2 are illustrated in Figure 1. It is easy to show that the bid price curve must have positive slope.² Lower bid price curves represent higher levels of utility.

The Opportunity Set

For given preferences, as represented by the family of bid price curves, choice of residence will depend on the opportunities that are available. The opportunities available will depend, in general, on the cost of providing public services, the assignment of households to jurisdictions, the price of land before tax in each jurisdiction and the local political decision-making process.

Suppose that the cost, G^j , of providing

the public service is an increasing function of the quality of service, g^j , and the number of households residing in the jurisdiction, N^j :³

$$(3) \quad G^j = G(g^j, N^j)$$

All jurisdictions are assumed to have the same cost function. For purposes of this paper, the cost function is assumed to exhibit "constant returns to scale":

$$(4) \quad G^j = \bar{G}(g^j)N^j$$

As noted earlier, local tax revenue, T^j , is raised through a proportional tax, τ^j , on the value of land:

$$(5) \quad T^j = \tau^j \tilde{p}_a^j D^j$$

where D^j is the amount of residential land in the j th jurisdiction. If each local government balances its budget, then:

$$(6) \quad \bar{G}(g^j)N^j = \tau^j \tilde{p}_a^j D^j$$

Since $\tilde{p}_a^j \equiv p_a^j(1 + \tau^j)$ and defining $n^j \equiv N^j/D^j$

$$(7) \quad \tilde{p}_a^j = p_a^j + n^j \bar{G}(g^j)$$

Given the population density (n^j) and the price of land (p_a^j), equation (7) represents the locus of opportunities available to the j th jurisdiction. Subject to this constraint, the vector (\tilde{p}_a^j, g^j) which is made available will be determined by the local political process.

Stratification by Wealth

Formulation of the equations of general equilibrium for this model would be a formidable task, but the problem becomes quite manageable if we assume that communities stratify perfectly by wealth. In particular, if all households have the same

² Apply the implicit function theorem to Equation (2) and observe that $\partial \phi^i / \partial g > 0$ and $\partial \phi^i / \partial \tilde{p}_a < 0$.

³ The type of public good assumed in this analysis has very special characteristics: quality enters the household utility functions as a "pure public good," but the public service is available only to residents and it is subject to "crowding" (cost is an increasing function of N^j).

utility function, then—to the extent that “voting with one’s feet” resembles a market process—it is natural to assume that households will sort out by wealth.⁴ However, it is not at all obvious that households will, in fact, stratify perfectly by wealth: the stability of perfect stratification hinges crucially on the nature of household preferences and the method by which local public services are financed.

In the absence of collusion, perfect stratification will be stable only if, once communities become stratified, no household decides to move to another jurisdiction. Figure 1 illustrates the case of a metropolitan area with two wealth classes (with $w_1 < w_2$) and two communities providing the packages represented by points A and B. If bid price curves have the shape portrayed in Figure 1, then stratification will be stable (in the sense defined above): households with wealth w_1 prefer A while households with wealth w_2 prefer B. Although a condition for stratification for the general class of utility functions can be derived [1, Ch. 3] a version of this condition more amenable to interpretation can be obtained if we assume that households have a nested CES utility function. In that case, stratification will occur only if public services are sufficiently complementary to land and the composite commodity. If the high public service jurisdictions are less accessible to jobs, the complementarity necessary for stratification will be somewhat less. If taxes are raised through a flat tax on households

⁴ Thus, in the case of two communities, all households with wealth less than w^* will reside in one jurisdiction while all households with wealth greater than w^* will reside in the other. General equilibrium can then be represented by a system of seven equations in seven unknowns ($p_d, p_g^2, \tau^1, \tau^2, g^1, g^2, w^*$). In this model, household preferences for the public service will be single-peaked which permits a simple formulation of the political decision-making process under majority rule. Extending this model to the case where there is more than one type of utility function is quite straightforward. (See Ellickson [1]).

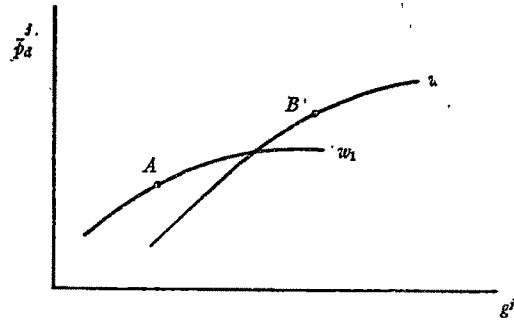


FIGURE 1

rather than a property tax, then the necessary condition for stratification will be satisfied in all but the case where the composite commodity and land are perfect substitutes for the public service.⁵

If collusion occurs, then stratification may be stable even though the conditions outlined above are not satisfied. Specifically, households in a wealthy jurisdiction may face a “prisoners’ dilemma”: an individual sale to a less wealthy household would be profitable but, if all residents engaged in such sales, the benefits of stratification would be lost. In this case, the residents of a wealthy community may restrict market operation through exclusionary zoning.

The Evidence of Stratification

There is some evidence that urbanization does produce stratification of political jurisdictions by wealth. In the classic *Suburbia*, Wood [6] concluded, after a survey of the existing literature, that few suburbs exhibit the heterogeneity characteristic of the metropolitan population as a whole: while some suburbs are very wealthy, others are extremely poor and

⁵ See Ellickson [1, Ch. 3]. The CES utility function takes the form $\{[a_1g]^{1/\rho} + [(a_2c)^{1/\omega} + (a_3d)^{1/\omega}]^{\omega/\rho}\}^\rho$ where a_1, a_2 and a_3 are constants and ω and ρ are substitution parameters. $1/\rho = 1 - 1/\sigma$ where σ is the elasticity of substitution, as usually defined, between g and the combination of c and d represented in the nested CES function. The necessary condition for perfect stratification is that $\sigma < 1$.

2. Oates, Wallace E., "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis," *Journal of Political Economy*, November-December, 1969, pp. 957-971.
3. Rothenberg, Jerome, "Local Decentralization and the Theory of Optimal Government," unpublished paper presented at the NBER Conference on Economics of Public Output, April 26-27, 1968.
4. Tiebout, Charles, "A Pure Theory of Local Expenditures," *Journal of Political Economy*, October, 1959, pp. 416-424.
5. Williams, Oliver P., Harold Herman, Charles S. Liebman and Thomas R. Dye, *Suburban Differences and Metropolitan Policies*, University of Pennsylvania Press, 1965.
6. Wood, Robert C., *Suburbia*, Houghton Mifflin Company, 1958.

Market Choices and Optimum City Size

By EDWIN S. MILLS and DAVID M. DE FERRANTI*
Princeton University

It is a safe guess that a survey would show a majority of social scientists, and a majority of economists, believe that a laissez faire economy produces excessive concentration of people and economic activity in a small number of large urban areas. Certainly, it is regarded as axiomatic in some quarters that the U. S. population problem is not excessive numbers of people, but their excessive concentration. A recent and representative statement is by James Sundquist [5].

Yet, economists have hardly begun to subject the problem to serious analysis. To do so it is necessary to start with some basic notions as to what cities are for. Two complementary ideas dominate the literature. First, scale and other agglomeration economies make it advantageous to concentrate production in certain central places where goods and services are produced for their inhabitants and those of the surrounding country-side. Lösch is of course the father of central-place theory. Second, regional comparative advantage makes it economical for each region to specialize in production to some extent and to trade with other regions of the country. It is then advantageous to locate the production of a region's exports and related goods at ports, railheads, and highways where interregional transportation costs are low.

Within the framework of the previous paragraph, there is no general theorem that the resulting urban areas are too large. In fact, Martin Beckmann [1] has shown in a pure central-place framework that competition may produce excessively

small urban areas. Of course, those who claim that some urban areas are too big have in mind considerations that are excluded from most central-place models. Pollution and congestion are most frequently identified as culprits. On the most abstract level, both can be subsumed under a general divergence between private and social costs as a function of city size. In an important paper [6], George Tolley has shown that such externalities do indeed produce cities of inappropriate size, but that they may be too large or too small depending on specific details of the model.

We do not have a detailed analysis of all the important issues related to optimum city size. In the next section we present a classification of alleged causes of excessive size, and some brief comments on each. The rest of the paper is devoted to an analysis of congestion, which seems to us to be the most important and interesting of the alleged causes of excessive city size.

I. A Classification of Possible Causes of Excessive City Size

An extensive sampling of the urban economics literature suggests that most of the factors that can plausibly be claimed to distort city sizes can be classified under three headings: the provision and financing of local public services, pollution, and congestion.

The focus of this paper is on the private sector, and we will therefore make only the briefest comments regarding the public sector. The literature contains a variety of claims that the terms on which transfers or public services are made available may provide certain groups of people, such as those with low incomes, with excessive incentive to live in large urban areas. Most

* The research reported in this paper was supported by a grant to the senior author from Resources for the Future.

of the analysis is too fragmentary to permit evaluation. Some claims appear to be based implicitly on the assumption of scale or agglomeration economies in the administration of local government programs. If so, central-place theory would seem to be the appropriate model, and we have already stated that it leads to no presumptions of excessive city size. Other claims are related to assumed redistributive programs of local government. We conjecture that a careful analysis would conclude that, in a country of fragmented local governments and mobile residents, local governments cannot accomplish, and generally do not attempt, substantial income redistribution. In any case, the issue is probably more relevant to the distribution of urban population between central cities and suburbs than to the total population of an urban area.

The case against pollution as a distorter of city size goes as follows. Assume realistically that the damage per resident of an urban area from health, aesthetic and other effects of pollution is an increasing function of the number and density of people in the urban area. Then the increment in pollution damage resulting from an additional resident must exceed the average damage. But residents bear the average pollution damage, and people will therefore locate in the urban area as long as their benefits, minus other costs of doing so, exceed the average pollution costs. That is to say, the marginal resident in the urban area will ignore the excess of marginal over average pollution damage. Thus, it is concluded, too many people will come to the urban area and it will be more polluted than would be efficient.

It is true that the urban area should be smaller if population is the only change one is willing to contemplate to solve the pollution problem. In that case, each resident should be charged an annual residence fee equal to the excess of marginal over average pollution damage, and it would make

the city smaller. But population of the urban area is by no means the only variable that can be used to affect pollution. In addition, we can reduce pollution per capita by increased recycling and treatment of wastes. Suppose that, through a system of effluent fees or by other means, our urban areas were induced to undertake recycling and treatment of wastes that were efficient in the usual welfare sense. Given an efficient pattern of waste recycling and treatment at each population level, determine the urban area's population so that it, too, is efficient in the sense that the addition of another resident adds as much to benefits as to costs of the urban area's activities. Then the question is whether the efficient population of the efficient city will be greater than or less than the present population of the city.

The answer depends very much on the mix of components in an efficient pattern of waste handling. One component will be raw discharges to the environment, relying on its capacity to absorb wastes. To the extent that raw discharge is optimum, it militates toward low densities, because high densities generate volumes of wastes that exceed the environment's absorptive capacity in an affluent economy. But the other components of an efficient pattern of waste handling will be treatment and recycling. To the extent that they are optimum, it militates in favor of high densities. A large part of the cost of treatment and recycling of wastes is their collection and transportation costs. These costs are generally much lower in high than in low density areas. Thus, an efficient policy toward waste disposal would call for small concentrations of people only to the extent that efficiency calls for substantial discharges of raw wastes, and it is doubtful whether it is a significant component of a waste policy appropriate to our time.

The case against congestion as a distorter of city size is similar to that against pollution. An urban transportation network has

a certain capacity to provide transportation without congestion, similar to the capacity of the environment to absorb wastes. If the use of the transportation system exceeds a rather low level, congestion results, so that travel times and travel costs rise. Beyond the point at which travel cost per trip rises with the number of travelers, the marginal cost to all users of the transportation system exceeds the average. But people take account of average travel cost in making locational decisions. Therefore, in deciding to locate in the urban areas, people ignore the congestion costs they impose on others, and the result is more people and greater use of the transportation system than would be optimum. At least at the micro level, this argument has been worked out much more fully, by Walters [8] and others, than has the analogous argument regarding pollution. At the level of an entire urban area, if measures to affect the urban area's population are the only ones that are contemplated to affect congestion, then measures should indeed be adopted to reduce the urban area's population. But, as with pollution, there are many other things that can be done to improve the efficiency of an urban transportation system and it is by no means obvious that the optimum size of an efficient city would be smaller than the size produced by existing conditions.

The first step in making progress with this problem is to find out what an efficient urban transportation system would look like. When traffic engineers ask that question, they want to know amounts of specific modes to put in specific places. Our goal is more modest. We simply want to know certain basic characteristics of an efficient system, to provide insight into the distortions that may be entailed by existing systems.

The functions of an urban area listed above are performed by specialized firms and other institutions, thus requiring a

large number of exchanges of goods and services. Exchanges require movement of goods and people, and the most fundamental characterization of an urban area is as a place where exchanges are facilitated by proximate locations of producing and consuming activities. The most extensively studied urban exchange is the sale of labor services by households and its related commuting activity. But it is only one of many kinds of urban exchanges. Thus the efficiency of an urban transportation system is to be judged by the exchange costs it entails. Viewed in that light, congested city streets may be extremely efficient. Travel is much faster on the streets around Princeton than on those in Manhattan, but exchange costs are high for Princeton residents because many of the goods and services wanted are not available locally. In large urban areas, and especially downtown, markets enable economical exchange by extreme concentrations of economic activity, i.e., by the use of high ratios of capital and labor to land in production. Up to a point, the result is efficient because short distances entail exchange costs that are small enough, despite slow travel, to offset the resulting high land values.

If one could assume that land rents reflected the opportunity cost of using land for transportation, it would be straightforward to compute optimum congestion in urban areas. To be specific, suppose that the cost per mile of travel at some place u , $p(u)$, is given by a function similar to one used by Vickery [7] and others:

$$(1) \quad p(u) = \bar{p} + \rho_1 \left[\frac{T(u)}{L_2(u)} \right]^{\rho_2}$$

Here, \bar{p} is travel cost per mile in the absence of congestion and the last term assumes a constant elasticity of congestion cost per mile with respect to the ratio of travelers $T(u)$ to land used for transportation $L_2(u)$. Total transportation cost at u is $p(u)T(u)$ and in an optimum system the

reduction in transportation cost from an increase in $L_2(u)$ would just equal land rent $R(u)$. For a pair of places u_0 and u_1 , the optimality condition gives:

$$(2) \quad \left[\frac{T(u_0)/L_2(u_0)}{T(u_1)/L_2(u_1)} \right]^{\rho_2+1} = \frac{R(u_0)}{R(u_1)}.$$

If u_0 were downtown and u_1 were near the edge of a large urban area, the right side of (2) might be about 100. Vickery believes that ρ_2 is at least two. That value would imply a downtown value of $T(u)/L_2(u)$ almost five times its value at the edge of the urban area, i.e., five times as many cars per lane mile downtown as at the edge of the urban area. Presumably, actual congestion ratios are much higher.

The objection to calculations like that above is that congestion may distort land values so that they do not reflect opportunity cost. The implication is that market choices do not result in an efficient locational and therefore travel pattern. Then, evaluation of properties of an efficient transportation system must start with a general equilibrium model of an entire urban area. We conclude the paper with an illustrative model.

II. A Model of Efficient Congestion¹

The following is almost the simplest model within which questions about optimum congestion can be asked. Suppose N workers are employed in a CBD of radius ϵ miles. N and ϵ are exogenous. At each distance $u (> \epsilon)$ miles from the CBD center, $\theta (< 2\pi)$ radians of land are available to house and transport the workers to the CBD. Each worker requires one house and each house requires a fixed amount of land, so $N(u)$ workers live on the $L_1(u)$ units of land devoted to housing u miles from the CBD center, where:

$$(3) \quad N(u) = a_1 L_1(u).$$

¹ The model in this section was suggested by those in [2] and [4].

The $L_2(u)$ units of land at u not used for housing are used to transport workers to the CBD, so:

$$(4) \quad L_1(u) + L_2(u) = \theta u.$$

All workers travel to the CBD, so the number of travelers at u , $T(u)$, is:

$$(5) \quad T(u) = \int_u^{\bar{u}} N(u') du'$$

where \bar{u} is the radius of the urban area and is endogenous. A crucial assumption is that travel cost depends only on distance to the city center and not on direction. It is assumed that (1) represents the time and other resources absorbed by the transportation system. Whatever resources are required to produce housing, they do not depend on the land allocation, and so are ignored. The only other real cost of the urban area is that each unit of land in the urban area has an alternative nonurban use, say in agriculture, that yields a rental R_A . Thus, the cost of the urban area is:

$$(6) \quad \int_{\epsilon}^{\bar{u}} [p(u)T(u) + R_A \theta u] du.$$

An efficient land use pattern is one that minimizes (6) subject to (1), (3), (4), (5) and the nonnegativity of all the variables. It is a problem in the calculus of variations, with one variable endpoint, \bar{u} . It is clear from (3), (4) and (5) that:

$$(7) \quad T'(u) = a_1 L_2(u) - a_1 \theta u.$$

Using (7), we can eliminate $L_2(u)$ from (6), derive the Euler equation in $T(u)$, and substitute $L_2(u)$ back in again, giving:

$$(8) \quad \frac{d \left[\frac{T(u)}{L_2(u)} \right]}{du} = - \frac{a_1}{\rho_2} - \frac{a_1 \bar{p}}{\rho_2 \rho_1 (\rho_2 + 1)} \left[\frac{T(u)}{L_2(u)} \right]^{-\rho_2}$$

Some experimentation indicates that if both \bar{p} and R_A differ from zero the resulting differential equations for $T(u)$ and $L_2(u)$ are intractable. The special case $\bar{p}=0$ implies the extreme assumption that uncongested travel is costless. $R_A=0$ implies the equally extreme assumption that urban land use is costless. We have chosen to concentrate on the former case, primarily because we have been able to obtain more interesting results.

With $\bar{p}=0$, we find from (8) and the necessary transversality condition on the endpoint \bar{u} that:

$$(9) \quad \frac{T(u)}{L_2(u)} = \frac{a_1}{\rho_2} (H + \bar{u} - u),$$

where

$$H = \frac{\rho_2}{a_1} \left(\frac{R_A}{\rho_1 \rho_2} \right)^{1/\rho_2+1}$$

which shows that congestion falls off linearly with u and reaches zero at \bar{u} if and only if R_A is zero. Equations (7) and (9) yield a first order differential equation whose solution, given $T(\bar{u})=0$, is

$$(10) \quad T(u) = \frac{a_1 \theta}{(\rho_2 + 1) Q(u)^{\rho_2}} \left[u Q(u)^{\rho_2+1} - \bar{u} H^{\rho_2+1} + \frac{Q(u)^{\rho_2+2} - H^{\rho_2+2}}{\rho_2 + 2} \right]$$

where

$$Q(u) = H + \bar{u} - u.$$

Using $T(\epsilon)=N$, we can express \bar{u} entirely in terms of known parameters.

Further manipulation leads to the following story. For small cities (i.e., small N and small \bar{u}), $L_2(u) < 0$, $L_1(u) > 0$, $L_2''(u) < 0$ and $T''(u) < 0$. A larger N requires a larger \bar{u} and a larger $L_2(\epsilon)$. If N is large enough so that $L_2'(\epsilon) > \rho_2 L_1(\epsilon)$, then $L_2'(\epsilon) > 0$. $L_2(u)$ increases with u until $L_2(u)/L_1(u)$ falls to ρ_2 , and declines thereafter. As before, $L_2''(u)$ and $T''(u)$ remain strictly

negative throughout, and $L_1''(u)$ remains positive. For still larger N , it may happen, depending on the values of R_A and ϵ , that $L_2(u)/L_1(u)$ increases for small u , and $L_2''(u)$ may be positive near ϵ . But it may happen that all the land at ϵ is used for transportation at a value of N too small for $L_2(u)/L_1(u)$ to increase with u near ϵ . Once N is large enough so that $L_2(\epsilon) = \theta \epsilon$, optimum land allocation will determine a u^* , $\epsilon < u^* < \bar{u}$, such that all the land within the interval between ϵ and u^* is used for transportation. Beyond u^* , the land allocation equations derived above apply.

The most interesting question to ask about the model has to do with optimum city size. Clearly, total transportation plus land cost rises as N increases. But does total cost rise more or less than proportionately with N ? We have found that, for some sets of parameter values, total costs rise less than proportionately with N for small N , i.e., for \bar{u} close to ϵ . For all admissible parameter values, costs rise more than proportionately when N is sufficiently large that $L_2(\epsilon) = \theta \epsilon$. Somewhere between, there must be at least one turning point. But we are unable to say more.

We have now characterized an optimum city in extremely simple circumstances. The next step is to ask how well utility and profit maximizing decisions will approximate the optimum city. But the present model is not rich enough for that purpose. Once the public sector has laid out an optimum transportation system, private decisions cannot distort land use in our model. To introduce a nontrivial private sector in the model, one of two changes must be made. First, make the amount or location of employment endogenous; or, second, permit varying factor proportions in the production of housing. Either change greatly complicates the model. But it is hard to see how else one can begin to think about optimum city size.

REFERENCES

1. Beckmann, Martin, "Equilibrium Versus Optimum: Spacing of Firms and Patterns of Market Areas," unpublished.
2. Mills, Edwin S., "The Efficiency of Spatial Competition," *Papers and Proceedings of the Regional Science Association*, Vol. xxv, pp. 71-82.
3. Neutze, G., *Economic Policy and the Size of Cities*, Canberra: Australian National University, 1965.
4. Solow, Robert M., "Land Use in a Long Narrow City," unpublished.
5. Sundquist, James, "Where Shall They Live?," *The Public Interest*, No. 18, February 1970, pp. 88-100.
6. Tolley, George S., "The Welfare Economics of City Bigness," unpublished.
7. Vickery, William, "Pricing as a Tool in Coordination of Local Transportation," in *Transportation Economics*, National Bureau of Economic Research, 1965.
8. Walters, Alan, "The Theory and Measurement of Private and Social Cost of Highway Congestion," *Econometrics*, Vol. 29, October 1961, pp. 676-699.

Social Problems and the Urban Crisis: Can Public Policy Make a Difference?*

By WORTH BATEMAN and HAROLD M. HOCHMAN

The Urban Institute

Originally, we had planned to call this paper "Social Problems and the Urban Crisis: Reflections on the Unheavenly Prospect," as a bow in the direction of Professor Banfield and *The Unheavenly City*. [2] For some reason, John Kain objected to that. Nevertheless, because 1970 has been the Year of Banfield and Benign Neglect, it seems appropriate to begin with a few observations on the relationship of Banfield's analysis to ours.¹

Both for Banfield and ourselves, the term urban crisis connotes a set of unwanted social and economic conditions. As Banfield is the first to acknowledge, it is disagreeable to be confronted with undernourished and poorly educated children, dilapidated housing and abandoned neighborhoods, or rising unemployment and high rates of crime. However, while Banfield's analysis reflects concern with these conditions, it implies that the more affluent members of society can do little about them, for it sees the lower classes as enmeshed in a culture of poverty which virtually guarantees the failure of positive action programs. In a sense, Banfield's interpretation suggests that the middle and upper classes are also victims of the urban crisis; perceiving the conditions in which

the lower classes live, they wish to help, but every attempt is frustrated because the patient fails to respond to treatment. In these circumstances, Banfield's counsel of resignation is a comfort to the frustrated do-gooder who can thus continue to live comfortably, though social problems remain unchanged.

Basic to our analysis, as to Banfield's, is the presumption that the urban crisis derives from the dissatisfaction of the lower classes. This dissatisfaction is based on their perception that the conditions in which they live are unacceptable in relation to what *they* would like them to be. The problem thus posed is primarily an urban one for two reasons: (1) the poor have tended more and more to concentrate in urban areas, and (2) the disparities between income and wealth are much more obvious in urban areas where the very rich and the very poor live in physical proximity. If either of these conditions did not hold, there would be no urban crisis *per se*.

Primarily, our analysis differs from Banfield's in its explanation of *why* our society has not made more progress in dealing with the crisis in urban communities. For Banfield, the "present-orientedness" of the lower classes is responsible for the failure of public policy; for us, it is the political majority that has failed, by not adopting the essential reforms in social and economic institutions. If this majority could remedy its own shortsightedness, it would recognize that such reforms are in its own self-interest. In other words,

*We wish to thank Stuart Altman, Richard Bird, Peter Brown, Richard Burton, Grace Dawson, Frank de Leuw, Harvey A. Garn, William Hamburger, William Gorham, Joel Rosenblatt, Alice Rivlin, and Kathleen Sproul for substantive and editorial criticisms of earlier drafts, though we alone retain responsibility for our interpretation of the urban crisis.

¹For a more detailed evaluation of Banfield's thesis, see [1].

our belief is that responsibility for the urban crisis stems neither from the lower classes nor from market failure, but from the deficiencies of the process of governance itself and the systems of rights it sanctions.

I. Dimensions of the Urban Crisis

In our view, two characteristics of our society are the heart of the urban crisis. The first is a judgment, held by significant political minorities, that the present degree of inequality in the distribution of income, wealth, and social opportunity is unjust; the second is the spillover effects of such distributional inequality: crime, family disintegration, poor education and health services, and the deterioration of the physical environment. While in themselves these problems are not peculiar to cities, the spatial concentration of people and industry which characterizes urban areas makes their effects both more acute and more apparent and, at the same time, facilitates the political organization of those who are most dissatisfied. Taken alone, however, the existence of these problems in urban areas is not sufficient to justify a diagnosis of "urban crisis."²

What has given the current situation its crisis proportions is that governments at all levels—local, state, and federal—have been unable to achieve a rate of progress, in dealing with distributional inequalities and their negative spillovers, acceptable to the individuals and groups pressing for

social change. [6, 9, 11, 15, 16, 18, 19, 20] Discrimination in employment and housing persists; public education is in a shambles in many urban school systems; wide disparities are evident in the quality and availability of such basic public services as fire protection, police, and sanitation; the incidence of poverty is high; and for generations the distribution of income has remained much the same. Most of all, the critical divisions between rich and poor in our society have increasingly been drawn along racial lines, even though racial differences in some socioeconomic measures have narrowed. Thus, despite some well-intentioned efforts and some progress in absolute terms, the Kerner Commission's portrayal of our society as "separate but unequal" seems as accurate now as it did in 1966.

Three interacting reasons can be cited for the failure to deal decisively with these problems. First, population mobility has produced patterns of settlement in metropolitan areas (within and among jurisdictions) that are sharply stratified geographically and therefore, politically, along economic, social, and racial lines. Second, the existing distribution of public authority within the federal structure and, in particular, the division (among levels and units of government) of the fiscal power to tax and spend has obstructed political action. Finally, the political process itself has tended to cater to pluralistic and powerful special interests. [2]

Metropolitan areas continue to collect a growing proportion of the nation's population, income, and wealth. Rising incomes in these areas are associated with suburbanization, a lateral movement which has had a distinctive racial bias. Low-income households and racial and ethnic minorities are increasingly concentrated in the central cities, with the relatively more affluent tending to the suburbs. On the one hand, this location process has increased

² Thus, as we use it, the term "urban crisis" connotes something more than the fact that urban society is not a utopia, that cities are not perfect and some wants of city dwellers go unsatisfied, as they always must, because public means are constrained. Not all externalities with which cities must deal originate in distributive injustice. Simpler externalities in consumption give rise to many urban allocation problems and are, in large measure, the basis for the public provision and financing of many local public services. But such problems, in our definition, are something different from the urban crisis itself and relate to it only to the extent that they interact with and exacerbate the primary causes identified in the text.

the extent to which central city residents must bear the fiscal and administrative burdens of providing basic public services; on the other, it has severely undercut the financial capabilities of central city governments. Central cities, without significant change in the structure of metropolitan government and its financing, can do little to deal with the distributional problems that pervade urban society and give rise to the urban crisis. Moreover, by accelerating the migration of industry and middle and upper-income residents to the suburbs, attempts to implement a strongly redistributive central city budget or to use public expenditure patterns in such areas as education and recreation as *de facto* means of redistribution are likely to be counterproductive. [3]³

While cities lack the resources and authority to deal effectively with the problems which confront them, other levels of government—state and federal—are limited in what they do by the political forces they represent.⁴ These forces reflect the interests of political majorities (or majority coalitions of minorities) who reside (and vote) primarily *outside* the jurisdiction of the city. The interests of these majorities and the legislators who represent them are frequently in conflict with the fundamental structural reforms needed to deal with distributive injustice and its external effects. This is seen, for example, in the actions of state legislatures in which rural-suburban coalitions block reforms intended to ease distributional inequities in school finance, law enforcement, and health services. [5] At the federal level, also, the Congress has been overwhelmingly effective in frustrating attempts to

reconcile the distributional problems of urban areas and overwhelmingly ineffective in adapting its procedures to the demands which an increasingly diverse and troubled society place upon it. [7] In part, as Banfield's perceptive critique of current public policy points out, this is because reasonable-sounding programs have been misguided. But much of the responsibility rests in the hands of vested interests—sectional, economic, and political—operating through the Congress or the President, opposed to distributional change or bent on diverting public resources to parochial goals, through such measures as oil depletion allowances, agriculture and maritime subsidies, educational aid to federally impacted areas, the interstate highway program, and, certainly not least, many military programs.

This situation is not new. Cities have always spawned more problems than their financial resources or political authority would permit them to resolve, and state and federal governments have always been heavily influenced by interests that appeared indifferent or even hostile to the people and problems of the city. If anything, the situation in an absolute sense may now be better than before, because cities, the disadvantaged, and racial minorities have benefited from such measures as voting rights legislation and reapportionment. Nevertheless, the situation in which the cities now find themselves is distinguished from the past by a number of factors, all of which justify a current diagnosis of "urban crisis." Our generally affluent society has heightened aspirations of all social and ethnic groups in relation to realizations, and the twin processes of urbanization and suburbanization have themselves widened the gap between actual and minimally acceptable social conditions. A series of constitutional and legislative victories have legitimized the grievances and claims for justice of racial and economic minorities. Thus encour-

³ What we refer to here—"the prisoner's dilemma" in which cities and suburbs find themselves—is well known. For a somewhat different diagnosis of the urban crisis and what to do about it, see [10].

⁴ Of course, these same political forces operate among districts and neighborhoods within central cities, but within cities the critical factor of rigid political boundaries is absent.

aged, these groups have stepped up their pressure for reforms to make such legal and moral rights a reality. However, the slow adjustment of social institutions and practices, plus the inevitable effects of a heritage of discrimination and prejudice on human attributes and capabilities, have kept measurable progress well below the pace of expectations.⁵

Thus, perceived distributive injustice, coupled with the urban concentration of those who are most dissatisfied, is what the urban crisis is all about. What triggers the crisis is a judgment that the structure of rights to property, income, and political power, and the institutions and rules determining the way in which these rights are established and enforced are inequitable and unlikely to change. [6, 8, 9] Repeated frustration of attempts to change prevailing practices increases the likelihood of recourse to violence or other extra-legal protest and separation (effective secession of the disaffected minorities) from the rest of society.⁶ In the extreme, neither of these alternatives is likely to be successful. The threat or fact of violence to protest perceived injustice makes more likely the formation of majorities which will enforce restraints on individual freedom, particularly the freedom of those most dissatisfied with the *status quo*. This multiplies the resentment, discontent, disagreement, disorder, and repression which is already characteristic of the urban scene. Extreme separation is also not a tenable solution, though it may be advan-

tageous for a small number of the disaffected. Unless the separatist minorities are willing to accept a permanent position of inequality (and inferiority) conflict will result. Thus, separation, as a long-term arrangement, seems unstable and unsustainable. [13] Moreover, like other social institutions which impede resource mobility, separation is inefficient. On balance, it will reduce not only the welfare of the political majority, but also that of minorities which practice it.

II. *Can Public Policy Make a Difference?*

Our diagnosis, if correct, suggests that the urban crisis is a crisis of legitimacy, deriving: (1) from a failure of existing institutions to produce outcomes consistently which all members of society can believe to be in their long-run interest, and (2) from a conviction that the process by which these outcomes are achieved is itself unfair. In our view, resolution of the crisis requires fundamental structural reforms in economic, social, and political institutions. [11] The accomplishment of this objective requires that the *de facto* system of legal, political, and economic claims be acceptable to all members of society. Such consensus must develop from a conviction that the obligations and rewards of a just and stable society are a matter of universal self-interest.⁷

If self-interest is narrow in scope, then the best to be hoped for is action to reduce the unwanted spillover effects of distributional inequalities that are felt by both majorities and minorities (either because of utility interdependence or because they affect individual welfare narrowly conceived). Stricter law enforcement, urban

⁵ For an excellent analysis of this point, see [14].

⁶ One side-effect of this situation is a loss of faith in the idea, basic to a free society, that rational and unfettered inquiry, given enough time, will yield a just and acceptable solution. One might argue that disadvantaged minorities themselves have never harbored any such presumption; however, it has surely been a basic premise of idealists, including students and intellectuals. To the extent that the urban crisis is manifested in violence and disdain for law, not just on the part of the disadvantaged but among some idealists at all levels of economic well-being, this belief appears to have broken down.

⁷ Little can be gained by arguing over a definition of consensus. Surely, it seems reasonable to assume that malevolent people are not a politically relevant force. If they were, a free society could not endure. In any case, true malevolence is an uncommon condition. Other things equal, few men prefer the Pareto nonoptimal to the good of one and all.

renewal, and a number of environmental programs fit this description. However, such measures are insufficient to set in motion the corrective forces needed to relieve the urban crisis because they do not come to grips with fundamental deficiencies in the "rules of the game" or "effective constitution" which govern social interaction and determine the distribution of income, wealth, and social opportunity. Use of the fiscal structure to redistribute income and wealth directly (e.g., universal income supplementation) is one way of effecting such change. In addition, other changes which govern market or nonmarket rules of the game (e.g., voting rights reform and open housing) can in due time significantly alter the distribution of private and social claims. The latter changes, in contrast to direct action programs, need neither strain the public budget nor require periodic authorizations, or annual appropriations, and do not rely on the bureaucratic structure for success.

At a practical level, all of this suggests that civil rights legislation and enforcement, income maintenance, open housing, and programs which assure equal access to education and equal protection of the law hold much more promise for the future of a "free" and harmonious (though still heterogeneous) urban society than an admixture of functional and categorical programs of limited scope. Urban renewal, the hodgepodge of *ad hoc* welfare programs now in force, and large-scale efforts to induce technological change and innovation in such areas as home construction and urban transportation are unlikely, or less likely, to make much of a dent in the urban crisis.⁸

⁸ Indeed, it is not always clear that disadvantaged families are the primary beneficiaries of programs ostensibly designed to help them or that the programs benefit the disadvantaged at all. One result of urban renewal, after all, has been neighborhood dissolution; another has been the opportunity for gain it presented

Changes in "rules of the game" which are basic to the urban crisis directly modify the effective structure of property rights and the structure of claims associated with them. In conjunction with market forces, such rights and claims largely determine the long-run distributional characteristics of our society and whether they are perceived as fair or unfair.

Thus, the system of private rights to property is very much bound up with the "urban crisis," not just because the distribution of such rights is unequal and the use of resources for private gain generates some diseconomies, but because these rights are not absolute. One man's rights are often another's deprivation. The issue of fairness or distributive justice arises, thus, in any discussion of property rights because the distribution of claims, which in a market-oriented society imply political as well as economic power, derives from such rights.

For a social practice like rights to private property to be acceptable to all members of a community, Rawls has argued it must be fair, and to be fair, it must satisfy two conditions: (1) all parties whom the practice will affect must be given an opportunity to make claims on its design,⁹ and (2) the practice must apply equally to all affected by it unless a deviation from equal application will benefit all concerned (e.g., unless a deviation is Pareto optimal, as might be the case if equal application, through effects on productivity, implies lower though more equal incomes for all.) [11] This definition provides a

to home builders and land developers [8]. Similarly, claims on the agricultural surplus are surely preferable to hunger. But are the subsidies which generate this surplus preferable to the direct income transfers which they might otherwise finance?

⁹ Note that this criterion provides a philosophical basis for black Americans, whose ancestors did not voluntarily opt into our society, to object to its existing system of property rights and prevailing distributive characteristics.

rationale for attenuating or modifying some existing property rights through a consent mechanism (e.g., through redistributive taxation, which alters rights to earned income). It also gives the reason why society characteristically proceeds with caution in doing so. The interests of those adversely affected are crucial in determining the degree and speed with which unjust practices are changed. It is one thing to begin *de novo* to build a system of law which restricts the rights attaching to private property; it is quite another to modify existing rights, thereby changing rules of the game in midstream.

The distinction we are making here is between "payoff problems," [12, 17] which are evaluated within a given system of rules, and "choice of game" problems, in which the rules themselves are decided. There are two ways of viewing the proper procedures for changing an unfair practice. They differ fundamentally on the issue of compensation. One approach holds to the view that the process of change requires compensation of the injured party; i.e., even the party now benefiting from an unfair practice whose interests will be adversely affected by a change in the practice. The second approach distinguishes between those claims which are legitimate and those which are not and holds to the view that the claims of those who benefit from an unfair practice can be overridden. These are primarily differences in moral views. As a practical matter, compensation or long lead times, or both, are generally associated with fundamental changes in the rules of the game.¹⁰

The problem of fairness in transition is, of course, less significant when the effect

of rule changes on individuals is indirect, as with congressional reform and measures designed to make federalism more effective (such as revenue sharing or changes in metropolitan or regional political structures). [5] It is more significant (and difficult) when rule changes affect individuals directly (as when property rights are modified or attenuated, since property and income attach to particular individuals). When rule changes *do* affect property rights directly, the assurance of orderly transition is greater if changes are such that they can work themselves out through market transactions. To be specific, open housing, which is a change in market practices, is clearly preferable to a system of residential quotas or enforced and uncompensated scattered-site housing.

An instructive approach to the dilemma of constitutional change—both in terms of the goals of change and the process for getting there—has been suggested by Buchanan. [4]¹¹ In his view consensus on basic changes in social institutions is more likely to be achieved if individuals have a dispassionate and detached view of their expected effects (because projection forward in time narrows the set of Pareto-optimal outcomes). This is true if constitutional-type changes, once adopted, do not take effect until a later point in time (say, 25 years), thus exempting the present generation from their effects. It may well be that we can do little to resolve the urban crisis equitably in the short run, but that in the long run, appropriate constitutional-type changes, capable of consensus, can bring about social justice and, thereby, dissipate the crisis itself. But, whatever its conceptual appeal, a delayed

¹⁰ Thus, for example, a community of risk-averse individuals, totally ignorant of their future income prospects, might choose a fiscal constitution which calls for equalizing redistribution. But immediate income prospects are not at all uncertain in the real world and consensus on such a rule is hardly realistic. [4]

¹¹ Buchanan has argued, as we have, "in fiscal theory, as in politics generally, scholars need to pay more attention to the working out of rules or institutions through which final outcomes emerge and less attention to the shape of these outcomes themselves," [4, p. 300]

solution (and much less one that puts change off for a full generation) does not come to grips with the interim problem of preserving a free society in the face of what appears to be a degenerative (perhaps rapid) movement away from consensus.¹²

III. Conclusion

Reduced to its common denominator, our discussion suggests that a "solution" of the urban crisis, in the short or long run, requires that the various constituencies in our nation be able to see the kinds of constitutional-type changes described (including changes in the fiscal constitution) as a matter of their own long-run self-interest. To attain justice with fairness, what is needed is a broad moral consensus, founded in a heightened perception of human interdependence and an understanding that a just and free society are the stakes of a game which must be played out within a viable time frame.

While we interpret the urban crisis quite differently from Banfield, some of our policy recommendations are similar to his. But while Banfield, in large part, sees the crisis as deriving from the present-orientedness of the lower classes, our analysis suggests that it is precisely the lower classes' ability to project an intolerable situation forward, together with the present-orientedness of the political majority and the political incentives of their representatives, which leads to action or inac-

tion that exacerbates or perpetuates the crisis.

Thus, while Banfield's counsel is one of pessimism and resignation, ours is not. We believe that resolution of the crisis is possible if political majorities are future-oriented enough to adopt constitutional reforms which not only benefit the lower classes but serve the majority's long-run self-interest. If these political majorities have the foresight to adopt fundamental, constitutional-type change, fulfillment can be harnessed to hope, and an urban society that is just, humane, and truly free can be a reality.

REFERENCES

1. Harvey A. Averch and Robert A. Levine, "Two Models of the Urban Crisis: An Analytical Essay on Banfield and Forester" (RAND, RM-6366-RC, September 1970).
2. Edward C. Banfield, *The Unheavenly City* (Little, Brown, 1970).
3. David F. Bradford and Harry H. Kelejian, "An Econometric Model of the Flight to the Suburbs," Econometric Research Program, Research Memorandum No. 116 (Princeton University, October 1970).
4. James M. Buchanan, *Public Finance in Democratic Process* (University of North Carolina Press, 1967).
5. Richard Burton, "The Metropolitan State: A Prescription for the Urban Crisis and the Preservation of Polycentrism in Metropolitan Society," presented before the Subcommittee on Urban Affairs, JEC, Oct. 15, 1970.
6. Eldridge Cleaver, *Soul on Ice* (Delta Publishing Co., 1968).
7. *Making Congress More Effective*, a Statement by the CED Research and Policy Committee (Committee for Economic Development, Sept. 1970).
8. Otto A. Davis and Andrew B. Winston, "Economics of Urban Renewal," *Law and Contemporary Problems*, Vol. 26 (Winter 1961), pp. 105-117.

¹² The idea of postdating change is not fully consistent with the premise that all parties whom a social practice will affect must be given claims on its design. Pragmatically, however, this counterargument is less troublesome than our prior reservation, for it is not necessarily inappropriate to interpret the constitutional preferences of any given social or demographic cohort as a surrogate for the preferences of its children. The implication of Buchanan's argument, moreover, is that the realistic alternative to fairness in the future is not fairness at present, which may be unattainable in any case, but the perpetuation of legally-sanctioned inequities.

9. William O. Douglas, *Points of Rebellion* (Random House, 1970).
10. Jay Forester, *Urban Dynamics* (MIT Press, 1969).
11. John Gardner, "The Common Cause," mimeo, July 20, 1970).
12. Harvey Leibenstein, "Long Run Welfare Criteria," Julius Margolis, ed., in *The Public Economy of Urban Communities* (Resources for the Future, 1965).
13. Edward H. Levi, "The Crisis in the Nature of Law," *The Record* (March 1970), pp. 3-23.
14. Elliott Liebow, *Tally's Corner* (Little, Brown, 1967).
15. Richard P. Nathan, "Reforms of the New Federalism," *New York Times*, Oct. 10, 1970.
16. Richard M. Nixon, "A Call for Cooperation," Presidential Message to Congress (Sept. 1970).
17. John Rawls, "Justice as Fairness," *Philosophical Review*, Vol. 67 (April 1958), p. 83.
18. Charles A. Reich, *The Greening of America* (Random House, 1970).
19. Elliott L. Richardson, "Towards a Workable Federalism," Address Before the National Association of Counties, Atlanta, Georgia, July 27, 1970.
20. Sir Geoffrey Vickers, *Science and the Regulation of Society* (Columbia University, The Institute for the Study of Science in Human Affairs, 1970).
21. Robert Wood, *The Radner Lectures*, delivered at Columbia University, April 7, 8, 14, 1970.

DISCUSSION

ROLAND ARTLE: An average reader of books and papers in Urban Economics today is likely to emerge from the reading with a sense of frustration—with a feeling of ending up in a cul-de-sac. He reads so many words about the acute urban problems, about the urban crisis, but the words do not seem to lead anywhere. He follows the winding thoughts of the renowned analysts in the field, only to find that they use partial-equilibrium approaches to problems which by all appearances and flavors are general-equilibrium (or general-disequilibrium) in nature.

However, reading Ed Mills is an entirely different story. In a series of ingeniously and skillfully varied, and truly pioneering, papers he has opened up new vistas. Although one cannot properly describe his framework as a general-equilibrium format, Mills has come a very long way in that direction, in comparison with the standard literature on location and urban land use. The main thrust of his present paper (written together with D. de Ferranti) lies in its analysis of congestion cost in relation to city size. The authors show that the elasticity of urban costs (as they define costs) with respect to city size (defined as the number of workers in the CBD) is less than unity for small cities and greater than unity for sufficiently large cities.

Assuming the discussant's role as devil's advocate, I shall now make some critical observations about the Mills-Ferranti paper—first some specific comments about a few of the assumptions underlying their analysis, and then some more general comments relating to their criterion of optimality.

A weakness of the analysis is that N , the number of workers employed in the central city, is exogenously determined. This does not, of course, mean that N is fixed; it can be varied parametrically. However, convenience comes at a price. With N exogenous, it was apparently necessary to make E , the radius of CBD, exogenous. In turn, this means that within the central area of the city, transportation and congestion costs are zero, the opposite of a real-world city. Re-

lated to this, there is also a slight anomaly in the calculation of city optimum on the last page, where the numerator includes workers in the transportation industry (\bar{p}), whereas the denominator (N) does not. Further, I question the validity of the form of the basic transportation (and congestion) cost function (1). It would be more meaningful with an explicit lower bound for the density variable:

$$\frac{T(u)}{L_2(u)} \geq b(N)$$

or:

$$\rho_1 = 0 \quad \text{for } \frac{T(u)}{L_2(u)} < b(N)$$

As this suggests, the lower-bound value should ideally be a function of city size, N .

Equation (2) is derived on the premise that "in an optimum system (without congestion distortion) the reduction in transportation cost from an increase in $L_2(u)$ would just equal land rent $R(u)$." However, this result also seems to imply:

$$\frac{\partial T(u)}{\partial L_2(u)} = 0$$

More importantly, it excludes corner solutions. Is there a deceptive analogy with land use in agriculture lurking behind this kind of reasoning? In principle, movement of goods and people can occur both underground and above ground. To assume that it can only occur on the ground level does seem rather limiting. Looking at the problem from a different point of view, one can also say that in urban analysis (possibly in contrast with agricultural analysis), it is undesirable to use (transport-service and other) production functions, $f(x)$, with the property that

$$f(x) = 0 \quad \text{for any } x_i = 0$$

$$x = (x_1, x_2, \dots, x_i, \dots, x_n)$$

Presumably, no other production function having this undesirable property is better known than the Cobb-Douglas function.

Now, for some more general comments. It should first of all be stressed that the (im-

plied) optimum-city-size criterion used by the authors is that of myopic, *allocative efficiency* (cost minimization). Distributional considerations, which are stressed by both Ellickson and Bateman-Hochman, might lead to smaller optimum city size.

Mills and de Ferranti say that two complementary ideas dominate the literature as to "what cities are for," namely central-place theory and (interregional and international) trade theory. Whereas this (unfortunately) may be a fair summary of the literature, it represents nevertheless a serious omission. Central-place theory is a highly descriptive approach with very limited explanatory value as a theory. The second strand of ideas, international trade theory, is in all essentials a static theory; using it to comprehend city phenomena may inadvertently lead us away from another notion which may be just as important in our understanding of why cities exist and grow, namely the dynamic notion of learning and technical change. Indeed, and essentially basing it on this concept, I would venture to hypothesize that *optimum city size is greater in the developing countries than in the developed countries*. There is in my view yet another notion which is also indispensable in our attempts to understand city phenomena, namely the concept of collective services (or with its better known, but somewhat misleading, name: "public goods"). Indeed, the ancient cities *were* public goods: they were centers of assembly and protection. The walls that enclosed the towns were an important collective service for the inhabitants; it is significant that the word "town" originally meant "enclosure." Other components in the bundle of collective services were the temple, the market place and at least the exterior of the palace. As cities grew in density and size, the bundle of collective services also grew. Lack of space precludes me from elaborating further on the significance of notions of learning and technical change and of collective services for our understanding of cities. The main point I wish to make is this: whereas Mills and de Ferranti concentrate their analysis to the cost side, in the determination of optimum city size, we cannot mean-

ingfully exclude the benefit or revenue side—and including it means facing the well-known conceptual difficulties of dynamics and of collective services.

Both collective services, or public goods, and public bads (exemplified by congestion in Mills-Ferranti) are manifestations of indivisibilities in consumption. As the authors say in their introduction, there is a strong tendency today to point to the multitude of public bads in the large cities (and to emerging and ominous new ones, such as urban guerilla warfare) in condemning the cities as being excessively large. It behooves one, however, to remind himself that cities were always *both* public goods and public bads. How exceedingly harsh the urban condition was for the majority of city dwellers in centuries past! I can think of no better single illustration than to cite the mortality rate among infants in Stockholm in the middle of the eighteenth century. An average of 387 babies per 1,000 died before reaching the age of one. That was almost twice as high as the national average at the time. (Incidentally, it is also about 35 times as high as today's figure for Stockholm.) What was it now again that Gertrude Stein used to say? "A city is a public good, is a good, is a good, is a public bad, is a bad, is a bad."

STEPHEN P. DRESCH: In response to Professor Mill's closing comment, I would only say that it may indeed be some time before economists have anything sensible or significant to say about optimal urban size. With reference to Bateman-Hochman, my comments would be too extensive; I cannot in a word do them justice.

I must immediately qualify my remarks by admitting that I have not perused Ellickson's more extended discussion of his model, which undoubtedly meets many of my narrower, more quibbling objections. But I do not consider this personal misfortune terribly significant in terms of the broad thrust of my comments. My remarks will be brief because my primary objective is to suggest, not to develop in detail, alternative approaches to the analysis of urbanization and local fiscal processes.

In commenting upon this paper—fourteen

years after Tiebout's "Pure Theory of Local Expenditures"—it is depressing to consider the degree to which we are not yet capable of meaningfully integrating conceptions of "political" or collective processes and those processes determining and constraining the development of urban spatial structure. This disappointing performance within urban economics is a graphic portrayal of the inadequacies of both public finance and urban location theory. The relevant public finance theory is still, in the main, a body of closed-system analysis, and location theory is still a theory of almost undifferentiated planes.

A major source of our difficulties in both fields, and in their merging, may be that we try to achieve too much. Following the simple analytics of, e.g., Muth-Alonzo location theory, we strive for general equilibrium models of urban processes, when in fact urban phenomena represent disequilibrium processes. This is, of course, the distinction between comparative statics and dynamics. I simply suggest that the two may be very tenuously related. In addition, as social managers as well as asocial scientists, we have been concerned with the grand design, with the capacity for globally normative judgements. This concern with general equilibrium and optimality reflects to a high degree our common analytical preconceptions, which I would not suggest we should attenuate too lightly. It may also reflect the urban economist's concern for his standing with his more established colleagues. Whatever the source, in this period of "benign neglect" and more or less disguised unemployment in the economics profession, when neither our analytical skills nor our practical judgements seem to be in strong demand, it may not be inappropriate to reconsider our conceptions and objectives and to identify alternatives and opportunities.

These comments are not meant to disparage Mr. Ellickson's paper, which takes the essentials of the Tiebout-Breton-Buchanan hypothesis and places them in such elegantly simple and comprehensive terms. I would argue, however, that the analysis does not serve to bridge the gap which its author identifies in his opening paragraphs. Even after Ellickson, the relevance of urbanization to the nature of

local government is not clear. A major contribution of the paper, however, is that it does identify, implicitly or explicitly, its own inadequacies and our sources of failure.

First, it is not clear within the confines of the assumptions why an urban area even exists. In adopting the "extreme assumption that the location of households is not constrained by accessibility to jobs," or apparently by accessibility to anything else, there appears to be no nongovernmental rationale for urbanization. (Given the restricted form of the public-good production function, there does not appear to be even a governmental rationale.) The Ellickson model is not an urban model in terms of any of the interactions which Bateman and Hochman, Mills and de Ferranti, or I would make the defining characteristics. The model does not suggest how we should relate differential accessibility to different types of employment and urban services, different levels of taxes and public services, and greater or lesser externalities.

Secondly, the Ellickson paper provides us with a component of a general equilibrium model of urban location and jurisdictional fragmentation. The model is essentially static, as the discussion of stability conditions makes clear. Thus, it is a model most relevant to the creation *de novo* or the instantaneous equilibration of an urban spatial structure characterized by jurisdictional fragmentation. However, the hallmark of urban development, and a central fact underlying "urban crises," is that urban growth is an historical, incremental process. Thus, Ellickson sees the isolated town, a closed system, as polar to the fragmented metropolis, an open system. In fact, there is a developmental process by which urban areas evolve, albeit a process more or less disjointed. Corresponding to this evolution in spatial form and economic complexity is a process of development in forms of social, including but not restricted to governmental, organization, or to use Bateman-Hochman terms, in *de facto* constitutions. While economies of scale undoubtedly explain a part of the development of local government, more important, if somewhat overlapping, sources are the external diseconomies which, given significant inequality in the distribution of

wealth, accompany increased density. In the isolated and *homogeneous* or egalitarian community these externalities are the exception rather than the rule; economies of scale and external *economies*, not diseconomies, are the sole rationale for collective action in resource allocation.

In the isolated inegalitarian community sufficiently low density may minimize negative externalities. Such interdependencies as do influence local public action derive from the social values of the dominant class(es). Only in a community characterized by inequality in the distribution of income and wealth (*and* by high population densities) do negative externalities (including the agitation of the lower classes) in themselves lead to increases in the scope and level of local governmental activity. Here we have a situation in which externalities cannot be avoided spatially and in which it is undesirable to avoid them governmentally.

Urbanization influences this process only to the degree that it increases population densities. If we define urbanization as a function of the *level* of population and the *degree* of interaction (exchange, etc.), then the impact of urbanization depends on the state of urban technology, e.g., housing and transportation technologies. If, with the existing technology, a given level of population and interaction requires high density, then negative externalities imply high incentives for collective, redistributive action.

But urban technology is not static. In an as yet unpublished paper Harrison and Kain examine changes in incremental densities over time and discover remarkably consistent and explicable patterns. These observed trends toward lower density may explain recent patterns of suburban stratification by wealth and the consequent cessation of annexation and opposition to metropolitanization, coupled with a lack of subdivision of larger jurisdictions, better than the Ellickson model, which does not explain why high-wealth groups remain in the large heterogeneous central city. In my alternative model they remain because of the accessibility of central locations. Negative externalities, like congestion, reduce, but

do not negate, the implied location rents attaching to central location.

Finally, the Ellickson model ignores intra-jurisdictional processes which may modify or negate apparent patterns of distribution, altering locational advantages. That these processes are empirically significant has been shown by Katzman and others with reference to the Boston City school system. On an analytic level the work of Crecine and others would suggest similar intrajurisdictional processes.

Other modifications, extensions and lacunae of the Ellickson model could be commented upon, e.g., the reasonability of the required utility-function complementarities and the failure to incorporate any household characteristics other than wealth. But the essential question is how these processes are to be approached in analysis. Extensions, such as Ellickson's, of the Alonzo-cum-Tiebout type are one alternative. As a result of such analyses as the Harrison-Kain study and of urban simulation modeling work such as that in which I have been involved at the National Bureau of Economic Research, my own guess is that a more explicitly dynamic and incremental approach, one which takes account of the rigidities and inertias of urban spatial structure, will be more fruitful in the long run. Further, the alternatives which I foresee are not, I think, necessarily subject to simultaneous equations bias or nonidentification.

In closing it is not clear to me that the Rothenberg-Ellickson type of formal general equilibrium model builds significantly upon the more literary general equilibrium models contained in the previous work of Tiebout, Breton and Buchanan. Furthermore, I think that there do exist alternatives which, while perhaps not as aesthetic, may provide more insight into processes of urban growth and development, the ostensible topic of the present session.

ANTHONY H. PASCAL: Reading the very worthy paper presented by Bateman and Hochman reminds one of the man who consulted his pharmacist about a remedy for a chronic complaint. He was given a beautifully designed new plastic container which had a

very impressive label reporting on the remarkable properties of the elixir. Taking it home and sampling it, he found the medicine to be about what he had been used to taking. All of this disappointed him, albeit in a soothing and rather reassuring way.

Like the pharmacist, I think the authors are about right in their diagnosis but deficient in their prescription. As far as explaining why we are in such a difficult state in America today, I would add only one item to their well-articulated list. Beside the increasing geographical concentration of the disadvantaged, the rapid rise in expectations communication and contiguity bring, and structural problems in the organization of government, it seems to me we must append the growing and rather deserved cynicism people feel about the ability of governments to deliver the benefits they promise. This is particularly true with respect to those programs broadly defined, and execrably titled, as human investment. Jurisdictional fragmentation, bureaucratic inertia, and greedy special interests share responsibility for these failures, but so does our scandalously incomplete knowledge of input/output relations in these programs.

That this is so is reflected in the experience of other countries—a source of insights to which Bateman and Hochman make no reference. The Dutch, the British, and the Scandinavians have more egalitarian societies than we and a good deal more social consensus on the definition of deserving groups. Still, if their government officials and social scientists do not yet speak of crisis, they certainly acknowledge urgent problems.

On to the prescription. To dismiss additional heavy-handed, bureaucratic interventions into housing, education, and transportation as unpromising is becoming part of the conventional wisdom, at least among economists and their fellow-travelers like Moynihan and Banfield. The label on the Bateman/Hochman nostrum, on the other hand, stresses wholesale reform in the rules of the game so as to alter permanently the social, political, and economic institutions, including those like property rights, which, as we all know, have divine origins. But when we open the package what do we find? Civil rights enforcement, open occupancy laws, income

maintenance (what happened to vouchers for schooling?). A remedy is no less effective, we must assume, for having been long in existence—even in the face of some possible violation of the new truth-in-packaging laws.

One would have liked, though, to have seen some explanation of more fundamental avenues of potential change. If, as I agree, the current degree of inequality in condition is morally objectionable and functionally dangerous, and if such inequalities are particularly galling when based on luck rather than on effort and ability, why not turn to the obvious device of 100 percent death duties? Fiscal experts assure us that the effects on work incentives would be minor. There would appear to be sufficient legal talent available to write tax laws both fair and administratively feasible. The proceeds of the duties could go to the provision of a national program which would award an annuity to each citizen on his eighteenth birthday.

If, as I agree again, territorially-based legislatures are inherently inequitable, let's experiment with alternate forms of representation. (We ought to be bold enough to discuss such ideas at professional meetings—even if we never actually try them out in New Jersey.) In a country as wired for communication as ours, some parts of legislatures might perhaps represent explicitly communal, religious, ethnic, occupational, corporate, and even avocational (e.g., nature-, or culture-freaks, etc.) interests. We do so little thinking about such basic changes that the practical difficulties they might generate seem overwhelming. Thinking through the difficulties, however, might prove highly instructive, if not directly useful. What is not at all clear is the usefulness of recording our votes for proposals which are already in the slow course of implementation, or are mouldering on the shelf because of the rigidities of the existing political structure.

Let me interject here some specific comments on various points raised in the Bateman/Hochman paper:

- They fail to explain why disadvantaged central city dwellers do not attain the kind of special interest power that other minorities—farmers, merchant-ship operators, arms builders,

tea tasters—seem to have. If lack of effective political organization is the cause, the authors' advocacy of integration and market-using programs seem beside the point.

- Postponing the impact of wealth redistribution so that it constitutes a welfare burden on the next generation is not exactly equivalent to a free lunch, or at least should not be thought so. Even acknowledging the excesses of the conservation nuts, the analogy with the depletion of natural resources retains some validity in this connection.
- Exhortation to do better is usually necessary but is rarely sufficient to bring about change. If it were, the handicapped would already be hired but Smokey the Bear would be unemployed. Incentives and sanctions are often found useful.

Which brings me finally to my last point. The authors seem to be working with the implicit model of a society which could reach a secure and permanent stasis in justice, equity, and fecundity, given certain quite manageable changes in practice and perception. It sounds a lot like long-run equilibrium. The imperialism of the economic view is truly insidious. Not to interpose counsels of despair in place of the authors' counsels of perfection, but it is altogether likely that in its larger dimensions society is better characterized as Kondratieff-cyclical or even yin/yang. Liberty alternates with repression, social concern with smug indifference, and puritanism with hedonism. This does not mean that we should do nothing, but it does suggest that minor adjustments will not bring about fundamental solutions.

THE ALLOCATION OF SOCIAL RISK

Social Risk and Financial Markets*

By WILLIAM BRAINARD, *Yale University*
and
F. T. DOLBEAR, *Brandeis University*

For our discussion social risk will be defined as uncertainty in social endowment—uncertainty about the aggregate consumption opportunities of society. The consequence of a given amount of social risk for the welfare of the individuals comprising society depends crucially on the financial arrangements society makes for its distribution. If financial markets were perfect, the distribution of risk would be Pareto optimal. In that case, prices, or “risk premia,” would be associated with each competitive equilibrium allocation indicating the terms on which individuals would be willing to trade certain income for risky outcomes. The magnitude of the risk premia or discount associated with a particular allocation depends, of course, on both the amount of social risk, and the preferences of individuals (degree of risk aversion). In the first section of this paper we will discuss the relationship between social risk and risk premia under the assumption of perfect markets.

Knowledge of the quantity of social risk and the magnitudes of risk premia are directly useful for the evaluation of government investment projects. In fact, much of the disagreement of various authors about the appropriate discount rate to use on government investment can be traced to a disagreement as to the magnitude of social risk. Also an empirical estimate of the magnitude of social risk

would provide a reference point for evaluating the efficiency of actual financial markets in distributing risk. The second section of the paper contains a brief discussion of the relevance of our analysis to the discounting debate and in addition some casual empirical evidence on the magnitude of social risk.

The perfect markets assumed for the first two sections of this paper enable individuals to trade on any particular contingency—the opportunities provided by a complete set of contingent commodity markets. In the third section of the paper we discuss the extent to which the markets for composite commodities—stocks, bonds, labor, etc.—are able to supply such opportunities. Such an analysis is useful both in providing an insight into the reasons particular types of financial claims may exist in the real world and in suggesting areas in which particular types of risk-spreading assets are missing.

I. *Social Risk and Risk Premia*

In order to explore the relationship between social risk and risk premia it will be useful to employ the Arrow-Debreu “state preference” (SP) or contingent commodity model. In the SP model contracts for the transfer of a commodity specify not only its physical properties but also an event on which the transfer is conditional. Preferences, production, and endowments are then all specified in terms of such “contingent” commodities, and a market is created for each such commod-

* Footnotes in this article were deleted by the editor in order to save space. Copies of the complete paper are available from the authors.

ity. The endowments we ordinarily call factors (labor, education, capital, etc.) are treated as bundles of contingent commodities. In Figure 1, for example, it is assumed that there are two states of nature defined by whether event E does or does not occur. The point labeled U corresponds to one unit of unskilled labor and entitles the owner to $\$X_E^u$ in the event E and $\$X_{\bar{E}}^u$ if not event E (\bar{E}). Similarly points S and K define vectors of returns for a unit of skilled labor and capital respectively.

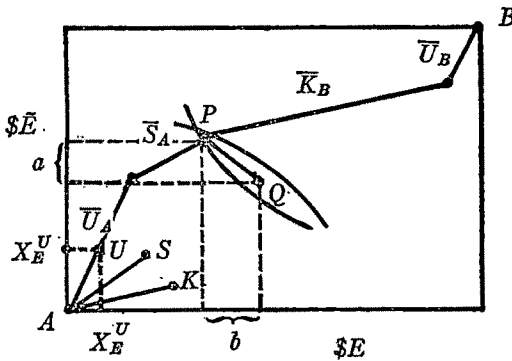


FIGURE 1

Differences in endowments and preferences imply that in general some trade will be necessary in order to achieve Pareto-optimal allocation. Consider, for example, the situation described in Figure 1, where, at the initial endowment point P , individual A has an endowment of $(\bar{U}_A, \bar{S}_A, 0)$ and individual B has an endowment of $(\bar{U}_B, 0, \bar{K}_B)$. The achievement of a Pareto-optimal point, such as Q , requires A to trade claims for $\$a$ in event \bar{E} , for $\$b$ in event E .

If individuals are vonNeumann-Morgenstern expected utility maximizers with diminishing marginal utility, i.e., risk averse, then it is a relatively trivial matter to demonstrate that in equilibrium prices will imply a kind of risk discount. To show this, suppose we have complete contingent commodity markets for S states

of the world. A typical individual who is endowed with \bar{X}_i in state i , will maximize:

$$E.U. = \sum_{i=1}^S \Pi_i u(X_i)$$

$$s.t. \sum_1^S P_i(\bar{X}_i - X_i) = 0 \quad X_i \geq 0$$

where $u'(\cdot) > 0$, $u''(\cdot) < 0$ and P_i is the price of the i^{th} contingent commodity. Then $\Pi_j u'(\hat{X}_j) = \lambda P_j$, $j = 1, \dots, S$.

For two different states of nature, k and 1,

$$\frac{\Pi_k}{P_k} u'(\hat{X}_k) = \frac{\Pi_l}{P_l} u'(\hat{X}_l)$$

So if $\Pi_k/P_k > \Pi_l/P_l$ then $u'(\hat{X}_k) < u'(\hat{X}_l)$ and obviously $\hat{X}_k > \hat{X}_l$. If individuals are agreed on probabilities, then in equilibrium, if we compare any two states, *every* individual will consume more in the state of nature with relatively high social endowment. It follows that in equilibrium the price of an expected dollar in a state of nature which has a large social endowment is lower than that of an expected dollar in a state with a small social endowment. In this model an individual may obtain \$1 for certain by purchasing one unit of each of the contingent claims at a total cost of $\sum P_i$. The amount by which the price per dollar of expected return (P_i/Π_i) for some contingent claim is less than the cost of \$1 for certain is what is usually called a risk discount. A positive (negative) discount simply reflects the fact that the marginal utility of income for a typical individual is lower (higher) than average in the particular state of nature in which the contingent claim pays off. Since the "price" of a composite commodity is the sum of the values of its component pieces, it is immediately apparent that the risk discount for a composite commodity is not a simple function of the variance in its possible outcomes. Roughly speaking, its discount will depend on

whether its returns tend to be in socially well or poorly endowed states of nature.

The particular market transactions an individual engages in, of course, depend upon his initial endowment. Although without knowing an individual's utility function we cannot precisely predict what his equilibrium commodity holdings or consumption will be, we have seen above that, if we order the states of nature by his consumption level, the order will be the same as if we order by social endowment. That is, he will consume most in the socially best endowed state, next most in the next best socially endowed state, etc. Hence we expect individuals to "insure" against contingencies in which they, but not society as a whole, stand to take large losses. The risk premium society requires to insure against such losses is less than any single individual would require (even if its returns are not negatively correlated with the returns on other assets). As a risk claim is divided up among an increasing number of individuals the risk premium a typical individual requires decreases more rapidly than the size of a typical share—in fact as the size of the share goes to zero, so does the premium required on all of the shares together. By spreading any particular risk claim across a large number of stockholders, an insurance company can reduce the total risk premium to a negligible amount, so long as its return is independent of other claims. Moreover, if an individual is fortunate enough to have an endowment which is negatively correlated with society's, he will find that the market will pay him a premium to "buy" his private risk. Such an individual will obviously be found "buying" stocks which are negatively correlated with his original endowment.

In the special case where there is no "social" risk in the sense that the social endowment in each state of nature is the same:

$$\sum_j \bar{X}_k^j = \sum_j \bar{X}_l^j$$

for all k, l , where \bar{X}_k^j is the endowment of individual j in state k . If all individuals are everywhere risk averse and agree on the probabilities, it is obvious that the only allocation which could be a competitive equilibrium would give a typical individual the same consumption in each state. Further, the equilibrium prices would be such that $\Pi_k/P_k = \Pi_l/P_l$ for all k, l , that is commodities would trade at their expected values.

As the preceding discussion of risk premia suggests, although it is theoretically possible to tell what expected return will be required privately and socially on any particular investment activity, in general it is not easy to summarize such requirements in terms of the price of a simple measure of "risk." In the special case where all individual utility functions are quadratic, however, it is a simple matter to find such a measure. For that case Sharpe, Lintner and others have shown that the risk premium on an asset depends on the covariance of its returns with the returns of a "market" index. It will be useful for our purposes to demonstrate this proposition in the context of a contingent commodity world.

Suppose that there are S possible states of the world, and contingent commodities exist for each state. One unit (\$1) of contingent commodity i yields R_i in state i and 0 otherwise. It will be convenient to regard an individual as having a choice among $S-1$ of these contingent commodities and a composite commodity called a "bond," rather than among the S contingent commodities. We will choose to eliminate the contingent commodity associated with the socially least well endowed state which may be taken to be state S . That is, $R_S \bar{\gamma}_S \leq R_i \bar{\gamma}_i$ $i = 1, \dots, S-1$ where $\bar{\gamma}_i$ is the social endowment of contingent commodity i . Since \$1 con-

tingent on state i can be purchased for $\$(1/R_i)$, it costs

$$\$ \sum_{i=1}^S \frac{1}{R_i}$$

to purchase $\$1$ for certain. Hence the gross rate of return on riskless bonds is

$$1/\Sigma \frac{1}{R_i} \equiv R_B.$$

Note that by creating a quantity of bonds (γ_B) which exhausts the supply of claims in state S , we reduce the supply of contingent commodities in other markets.

The problem for an individual with a quadratic utility function is:

$$\max U(\bar{R}, \sigma_R^2) \text{ s.t. } \sum_{i=1}^{S-1} \gamma_i + \gamma_B = \omega_0,$$

where $U_1(\cdot) > 0$, $U_2(\cdot) < 0$, where ω_0

is his initial wealth, γ_i is the amount invested in commodity i , γ_B is his holding of bonds, and where

$$\bar{R} = \sum_{i=1}^{S-1} \pi_i R_i \gamma_i + R_B \gamma_B$$

$$\sigma_R^2 = \sum_{i=1}^{S-1} \pi_i R_i^2 \gamma_i^2 - \left(\sum_{i=1}^{S-1} \pi_i R_i \gamma_i \right)^2$$

The first order conditions for a maximum are given by:

$$(1) \quad U_1(\cdot) \pi_i R_i + 2U_2(\cdot) \left[\pi_i R_i^2 \gamma_i - \pi_i R_i \sum_{k=1}^{S-1} \pi_k R_k \gamma_k \right] = \lambda;$$

$$i = 1, S-1$$

$$(2) \quad U_1(\cdot) R_B = \lambda$$

Hence

$$(3) \quad (\bar{R}_i - R_B)$$

$$= \frac{2U_2(\cdot)}{U_1(\cdot)} \left[\pi_i R_i^2 \gamma_i - \bar{R}_i \sum_k \pi_k R_k \gamma_k \right];$$

$$i = 1, S-1 \text{ where } \bar{R}_i = \pi_i R_i$$

It is apparent from (3) that the propor-

tions of contingent commodities held (γ_k/γ_i) are independent of the individual's preferences. Individuals who have quadratic utility functions and agree on the distribution of returns (probabilities) will hold the same relative proportions of the $S-1$ contingent commodities. All of the relevant opportunities for diversification can therefore be provided by riskless bonds and a mutual fund pooling the $S-1$ contingent commodities.

By multiplying each equation in (3) by $\gamma_i/\Sigma \gamma_i$ and summing we may find an expression for $-2U_2(\cdot)/U_1(\cdot)$ in terms of the expected return and variance of such a fund.

$$(4) \quad (\bar{R}_S - R_B)$$

$$= - \frac{2U_2}{U_1} \left[\frac{\pi_i R_i^2 \gamma_i^2}{\Sigma \gamma_i} - \bar{R}_S \Sigma \gamma_i \right]$$

$$= - \frac{2U_2}{U_1} [\sigma_S^2 \Sigma \gamma_i]$$

where \bar{R}_S and σ_S^2 are the mean and variance of return respectively on a mutual fund with proportions $\gamma_i/\Sigma \gamma_k$ of the $S-1$ contingent commodities. Substituting this into equation (3) we obtain

$$(5) \quad (\bar{R}_i - R_B)$$

$$= \frac{(\bar{R}_S - R_B)}{\sigma_S^2 \Sigma \gamma_i} (\pi_i R_i \gamma_i - \bar{R}_i \Sigma \pi_k R_k \gamma_k)$$

$$= \frac{(\bar{R}_S - R_B)}{\sigma_S} \left(\frac{cov_{i,S}}{\sigma_S} \right)$$

The coefficient $(\bar{R}_S - R_B)/\sigma_S$ may be termed the market price of risk and $cov_{i,S}/\sigma_S$ is the nondiversifiable risk (in Lintner's terminology) associated with a unit of the i^{th} asset.

From (5) it follows directly that the required rate of return on any composite security (c), is given by

$$(6) \quad (\bar{R}_c - R_B) = \frac{(\bar{R}_S - R_B)}{\sigma_S} (cov_{c,S}/\sigma_S)$$

II. *Risk Discounts and Government Investment*

As long as we accept an individualistic view of society and perfect markets exist, the risk premia discussed in the preceding section provide a means for "discounting" for risk on government investment. In fact, a number of authors have used the state-preference model as the basis of their policy recommendations. J. Hirshleifer (pp. 276-277), for example, concludes: "The efficient discount rate, assuming perfect markets, is the market rate implicit in the valuation of private assets whose returns are 'comparable' to the public investment in question—where 'comparable' means having the same proportionate time-state distribution of returns." On the other hand, Arrow and more recently, using the same model, Arrow-Lind (AL) have argued that the government should behave as an "expected value decision maker." This conclusion derives immediately from what AL call the "plausible" assumption that government investments are "small" and independent of other components of national income. This is, of course, precisely what is indicated for the quadratic case by equation (6) above. If an investment (G), government or private, is small and its returns are independent of other components of income, then $cov_{G,S}$, the covariance of its return with the index of returns on all assets including itself, will be small relative to σ_S —a negligible risk premium will be required.

The reason for AL's assumptions being plausible is hard to find in their article. They simply state (p. 373) "some government investments . . . are interdependent . . . (and) should be evaluated as a package. Even after such groupings . . . there will be a large number of essentially independent projects." AL make the same presumption of independence when they argue that if managers (p. 376) were "acting in the interest of the firm's share-

holders, they would essentially ignore risks and choose investments with the highest expected returns." If the assumption is correct for private as well as government investment (it is difficult to see why, if it is plausible for one, it isn't plausible for the other) then social risk must be zero. Baumol, following the same line of reasoning, comes to exactly this conclusion.

That social risk is zero is not an obvious fact. Although there is a shortage of empirical work bearing directly on the question, there is a good deal of evidence about the high correlation of returns within some sectors of the economy. Even the most casual examination of the stock market over the last few years reveals that society's valuation of the total claims on future consumption traded on the stock market are capable of wide and unpredictable fluctuations. Feeney and Hester, in a provocative study of the stock market, find that one can explain a very large percentage of the variation in return to any of the Dow Jones stocks with the same very small number of independent components. George Douglas obtained similar results for a large number of stocks traded on the N.Y.S.E. It seems quite possible that a small number of factors may be responsible for a large fraction of the variation of returns on most investments, with a resulting high correlation in their returns. In the U. S. economy, the business cycle itself is undoubtedly one of the major reasons for such correlation.

These results tend to support the view that the risk premia observed in the stock market reflect the presence of risk which cannot be eliminated by efficient distribution within the market itself. If other components of income have comparable risk associated with them and are positively correlated with corporate earnings, then there is a substantial amount of social risk. Although some investments, private or government, may be independent of, or even negatively correlated with, other

components of national income, and therefore deserving of zero (or negative) risk premia, many will require a positive premium. If, on the other hand, other components of income have less risk and/or are independent or negatively correlated with corporate earnings, social risk may be quite low.

The empirical problem of deciding what social risk actually is, and how much private risk could be reduced by pooling across various sectors of the economy, is a difficult one, and beyond the scope of this paper. Some suggestion of the magnitude of social risk and the possible gains that could be had if some way were found to efficiently distribute risk from all sectors can be gotten by looking at the variances and covariances of various components of national income. Table 1 shows the coefficient of variation (on the diagonal) of components of real national income and their correlations with each other, after removal of each component's trend.

In that case the coefficients on the diagonal give some indication of the magnitude of risk which would be borne by a typical individual in a particular group if markets existed for efficiently distributing risk within a group, but not across groups.

It is perhaps not too fanciful to think that such opportunities for efficient diversification exist for corporate profits. In that case, the difference between .09 and the contribution of corporate profits to social risk (which can be easily obtained by application of the Sharpe formula (Equation 6)) is a measure of the divergence between private and social risk. The low correlation of profits (.07) with national income indicates that the private risk borne by a typical owner of corporate profits is likely to greatly exceed the social risk on corporate profits. If one were more confident of the estimates of variances and correlations in Table 1, the risk premia which various authors have estimated could be used to value the gain from elimi-

TABLE 1—COEFFICIENTS OF VARIATION AND CORRELATION OF COMPONENTS OF NATIONAL INCOME—1947-1965

	(0)	(1)	(2)	(3)	(4)	(5)
(0) National income	.04					
(1) Compensation of employees	.15	.06				
(2) Business and prof. income	.44	.16	.02			
(3) Farm	-.04	-.21	-.22	.12		
(4) Rental income	.39	.19	.36	-.66	.08	
(5) Corporate profits	.07	.13	.54	.35	-.10	.09

As can be seen from the table, the uncertainty associated with most of the major components of income exceeds that on national income as a whole. This reflects the fact that although many of the components are positively correlated, the correlations in general are not large, and in some cases (notably those involving agriculture) are actually negative.

Suppose it is assumed that there is a rough correspondence between components and particular groups of individuals (laborers, entrepreneurs, capitalists, etc.).

nation of this discrepancy. In other sectors, a major gain would undoubtedly come from improvement in the opportunities for diversification within the sector. Table 1 suggests, however, that for other sectors as well (e.g., the farm sector), the gain from diversifying across sectors would be substantial.

III. *Financial Markets and Private Risk*

In the usual formulation of a contingent commodity model, an individual's factor

endowment is described as a vector of contingent commodities. A brain surgeon, a laborer, and a share of AT&T stock are simply different combinations of contingent commodities. The fact that a typical contingent factor endowment comes "packaged" as units of capital or labor is of no importance to an individual as long as a complete set of contingent commodity markets exist and, as is usually assumed, transactions are costless. An individual endowed with some composite factor could "slice off" and sell contingent shares. In Figure 1, for example, it is immaterial to individual *B* that his endowment of $(\$E, \$\bar{E})$ comes in the form of unskilled labor and capital.

In the real world there are costs, both in running markets, and in conducting transactions which are likely to make it uneconomic to have a complete set of contingent commodity markets. A major part of the explanation of the existence of many of the markets and institutions we actually observe is undoubtedly that they provide a means of economizing on such costs. In comparing the performance of a complete set of markets for contingent commodities (henceforth simply called *claims*) with some alternative set of markets, we obviously should take into account differences in these costs as well as differences in the opportunities for trade. In this section we will limit ourselves to a discussion of the extent to which markets in the composite commodities (e.g., "labor" or "securities," henceforth simply called *commodities*) provide all the opportunities required for the efficient distribution of risk. This question breaks naturally into two parts: (1) when will a set of commodity markets provide the same consumption opportunities as a complete set of claim markets, and (2) when can the trade required for efficient distribution take place without all the opportunities provided by a complete set of claim markets? It may not surprise the reader that the requirements for (1) are

quite stringent, and that the requirements for (2) are restrictions on the differences in preferences and/or initial endowments across individuals.

Suppose there are n commodities and S states of nature. One unit of a typical commodity (j) provides a return of δ_{ij} in state of nature i . An individual who purchases γ_j ($j=1, n$) units of the various commodities has effectively "purchased" $c_i = \sum_{j=1}^n \delta_{ij}\gamma_j$ units of contingent claim i . In matrix notation an individual's consumption possibility locus is given by:

$$(S \times 1) \quad (S \times n)(n \times 1) \\ c = D\gamma \quad \text{s.t.} \quad \sum_{i=1}^S P_i(\gamma_i - \bar{\gamma}_i) = 0$$

where P_i is the price, and $\bar{\gamma}_i$ the individual's initial endowment, of the i^{th} composite commodity. The matrix $D \equiv \{\delta_{ij}\}$ represents the linear transformation from commodity space to claim or state space. It is obvious that if the rank of D is S , then an individual who is allowed short sales can "purchase" claims in any proportions, and indeed can do so by engaging in trades in only S commodities. In Figure 1, for example, an individual can achieve any combination of contingent claims by trading in any two of the three composite commodities (unskilled labor, skilled labor, and capital). Hence, in this case, a complete set of "factor" markets provides all of the opportunities that a complete set of contingent claim markets would, including the possibility of making a riskless investment.

Under these assumptions the market equilibrium which would result from trade in the n commodity markets would be equivalent to the equilibrium which would result with a complete set of S claim markets. There would, of course, be a good deal of structure to the commodity prices in equilibrium. The price of any commodity which can be represented as a linear combination of some other set of commod-

new asset, or market, will be required to achieve Pareto optimality. If a riskless security does not exist, however, and there are any differences in the degree of risk aversion across individuals, then equilibrium in the commodities market will not be Pareto optimal.

Creation of Bonds

Under these assumptions, differences in risk aversion provide an incentive for firms or individuals to create a riskless asset (assuming one does not already exist). From the point of view of the model it does not matter who issues such "bonds," so long as they are forthcoming in sufficient supply to allow all necessary transactions between more or less risk averse individuals. One can imagine, for example, each firm "slicing" off equal amounts of return in each state of nature and packaging them as bonds. The maximum amount of bonds a particular firm could issue in these circumstances would be dictated by its least well-endowed state. The total number of bonds that could be created in this fashion might not be adequate to achieve Pareto optimality; it might be necessary to allow individuals or other financial institutions to issue bonds on the collateral of their holdings of risky securities. The maximum amount of bonds that might conceivably be required could be quite large—it might be necessary to create as many bonds as are allowed by the total returns to society in the least well endowed state of nature. This is the quantity we "created" in Section I, and would be sufficient to guarantee the possibility of Pareto optimality.

Allocation of Specific Risks

As mentioned above, there are significant costs which provide incentives to economize on the number of markets, and to select those which economize on the number of transactions required to achieve an efficient allocation of commodities. These incentives are important in under-

standing the institutions and markets society has created for diversifying different kinds of risks. For small (from the point of view of society) independent risks, efficient allocation requires that a large number of people buy pieces of the action. The need for this type of diversification does not depend on differences in tastes. The efficient allocation of claims which embody social risk, on the other hand, requires a mechanism for taking account of differences in the degree of risk aversion. The above analysis suggests that dividing the returns to capital between relatively riskless bonds and common stocks may indeed be an effective way of economizing on the number of markets required to achieve an efficient allocation of social risk.

Such arrangements do offer opportunities for the spreading of some kinds of risks. Unfortunately, not all risks seem well handled by existing financial markets and institutions. We would like to discuss two of these. First, consider the allocation of risks associated with labor income. Even though there are substantial private risks associated with the labor income of individuals, we do not observe markets for the purchase and sale of labor claims. As in the case of other risks, transactions costs are part of the reason for the absence of such markets. It is also true that individuals and unions in negotiating their labor contracts can "sell" variability in the returns to labor which firms can in turn sell in the stock market. In effect the firm can act as a broker for labor. This provides a mechanism for the efficient allocation of labor earnings without a recognizable market for labor claims. Given the fact that the returns to ownership of a given firm are likely to be positively correlated with the fortunes of its employees, the limits on this practice are similar to those which limit the issue of debt.

Although transactions costs might be important, probably the major reason why opportunities for the diversification of

labor income do not exist is the presence of "moral hazard." Moral hazard occurs whenever the probabilities associated with states of nature are not completely exogenous from the actions of the parties to a contract. More specifically, moral hazard exists if one of the parties might, without detection, violate the rules of the contract concerning his behavior, or misinterpret which state of nature has occurred. If a laborer is able to vary the extent to which he obeys the rules (e.g., casually execute his duties, making it more likely his income will be low) or control in some way the interpretation of the state of nature (e.g., falsely claim he doesn't have a job) and if it is in his interest to do so, there will be a moral hazard problem.

Are there any arrangements which could facilitate the spreading of risk on labor income and which would avoid the moral hazard problem? For many workers, a sizable component of the variability in income is beyond his control. For example, an individual economist may by his efforts affect his relative position in the profession and hence his income. However, most of us would agree that the rapid increase in our incomes during the late 1950's and most of the 1960's, as well as the slowdown in the last two years, had little to do with our own performances. The same thing goes for firemen, engineers, and many other labor groups. Thus it should be possible to sell a portion of such exogenous variability in income streams. In particular we would suggest there should be a contract called "economists' median income." An economist would agree to deliver (say) 25% of economists' median income as reported by the BLS in (say) 1985. He would take the proceeds from this sale and buy a portfolio of the future incomes of other labor groups. If economists' future salaries are not perfectly correlated with the future salaries of others, mutually beneficial risk-spreading transactions will be possible. Presumably, such securities

would be handled by brokers who would specialize in future income securities.

Finally we would like to discuss briefly opportunities for diversifying uncertainty in retirement income. Many retirement plans offer fixed annuities based on payments over the working lifetime, so there is no risk in dollar income for the retirement years. However, since the elderly do not eat dollars, many have lived to regret that they had no hedge against inflation. A partial answer to this problem has been provided by variable annuities where the level of payments is tied to earnings or prices of common stocks. But this is not a completely effective hedge since—as any observer of the current scene knows—stock prices don't always move with commodity prices. The security which would ameliorate such problems, of course, is the real income bond.

But not all risk-averse individuals are directly concerned with fluctuations in real income. Some may be interested in their relative positions. They would feel left behind if there were rapid increases in real income for the general public during the period they were on fixed real incomes. For such people we would propose relative income, or "Duesenberry" bonds. In their most imaginative form, an individual could name his reference group, though of course he would have to bear the expected value of any increased transactions costs incurred in keeping abreast of the real incomes of his individualized peer group.

REFERENCES

- Kenneth J. Arrow, "Discounting and Public Investment Criteria," in A. V. Kneese and S. C. Smith, eds., *Water Research* (Johns Hopkins Press, 1966), 13–32.
- Kenneth J. Arrow and Robert C. Lind, "Uncertainty and the Evaluation of Public Investment Decisions," *Amer. Econ. Rev.*, June 1970, 60, 364–378.
- William J. Baumol, "On the Social Rate of Discount," *Amer. Econ. Rev.*, Sept. 1968, 58, 788–802.

The Role of Insurance in the Allocation of Risk*

By RICHARD KIHLMSTROM

Northwestern University and the University of Massachusetts
and MARK PAULY, *Northwestern University*

The purpose of this paper is to develop a general equilibrium theory of insurance and to use it to study the allocation of risk. Our theory is based on a reinterpretation of the contingent claims models of Arrow [1], Debreu [7] and Radner [10]. As interpreted here the Radner model allows us to introduce the cost of the information required to provide insurance. This provides an explanation for the observation that real world insurance institutions provide less insurance than suggested by the Arrow-Debreu model.

In Section I, we reinterpret the Arrow-Debreu model to obtain a theory of insurance without information costs. In Section II, the Radner model is used to introduce information costs. In the remainder of Section II and in Sections III and IV we study the nature of Pareto-optimal insurance allocations when information is costly.

I

We begin by introducing a highly simplified model of an economy which is designed to focus attention on the questions associated with insurance.¹ There are n consumers, each of whom receives income in one physical commodity and consumes only that commodity, and there are T (finite) states of nature $\{1, \dots, T\}$. Following Arrow [1] and Debreu [7], we

consider consumption of income in state t as a different commodity from consumption of income in state t' . We let $I(t)$ be the amount of income in state t . We suppose that nature has endowed consumer i with a certain amount S_i of income, but that it also subjects him to losses $x_i(t)$, the size of which depend on the state of nature that prevails. Thus the initial allocation of consumer i is the vector $(S_i - x_i(1), \dots, S_i - x_i(T))$. For simplicity and convenience we also assume that $x_i(1) = 0$ for all consumers, and that $x_i(t) \geq 0$ for all i and t .

Each individual, i , has a probability distribution $Pr^i(\cdot)$ over the set $1, \dots, T$ of states, and his utility is given by

$$U_i(I_i(1), \dots, I_i(T)) = \sum_{t=1}^T u_i(I_i(t)) Pr^i(t),$$

where u_i is a risk averse von Neumann-Morgenstern utility function.

Suppose that in such an economy, individuals are able to buy and sell *insurance policies* $\alpha = (\alpha(1), \dots, \alpha(T))$. The policy α is one which pays $\alpha(t)$ units of income in state t and costs $P(\alpha)$ units of income. If individual i buys policy α_i his income in some state t is therefore

$$I_i(t) = S_i - x_i(t) + \alpha_i(t) - P(\alpha_i).$$

We can then define an insurance equilibrium as an insurance price function $\hat{P}(\alpha)$ and a vector of policies $\langle \alpha_1, \dots, \alpha_n \rangle$ such that α_i maximizes

$$\sum_{t=1}^T u_i(S_i - x_i(t) + \alpha_i(t) - \hat{P}(\alpha_i)) Pr^i(t),$$

and the following supply-demand condi-

* Research support from the National Science Foundation (for Professor Kihlstrom) and from the Health Services Research Center of Northwestern University and the American Hospital Association (for Professor Pauly) is gratefully acknowledged.

¹ Borch [5] calls this a "reinsurance market."

tions hold:

$$\sum_{t=1}^n \alpha_i(t) = 0 \quad \text{for all } i,$$

and

$$\sum_{i=1}^n \hat{P}(\alpha_i) = 0.$$

Note that two characteristics of the equilibrium price function are that $\hat{P}((1, \dots, 1)) = 1$, and $\hat{P}(a\alpha + b\beta) = a\hat{P}(\alpha) + b\hat{P}(\beta)$, for all policies α, β and positive numbers a, b .

The relationship between this concept of an insurance equilibrium and the concept of a contingent claims competitive equilibrium as defined by Arrow and Debreu should be clear. Let $I_i = \langle I_i(1), \dots, I_i(T) \rangle$. A competitive equilibrium is a vector of prices $\langle \hat{q}_1, \dots, \hat{q}_T \rangle$ and a vector of income vectors $\langle \hat{I}_1, \dots, \hat{I}_n \rangle$ such that \hat{I}_i maximizes

$$\sum_{t=1}^T u_i(I_i(t)) Pr^i(t)$$

subject to

$$\sum_{t=1}^T \hat{q}_t [I_i(t) - S_i + x_i(t)] = 0$$

and supply equals demand; i.e.,

$$\sum_{i=1}^n [\hat{I}_i(t) - S_i + x_i(t)] = 0$$

for all t .

For any income vector $\langle I_i(1), \dots, I_i(T) \rangle$ it is convenient to define the corresponding vector of losses $z_i = \langle z_i(1), \dots, z_i(T) \rangle$ borne by i by $I_i(t) \equiv S_i - z_i(t)$. Then a competitive equilibrium can be viewed as being an equilibrium vector of loss shares $\langle \hat{z}_1, \dots, \hat{z}_n \rangle$. The equilibrium insurance policy $\hat{\alpha}_i$ will be related to \hat{z}_i by

$$(1) \quad \hat{\alpha}_i(t) - \hat{\alpha}_i(1) = [x_i(t) - \hat{z}_i(t)] - [x_i(1) - \hat{z}_i(1)]$$

and the equilibrium price function \hat{P} will be related to $\hat{q}_1, \dots, \hat{q}_T$ by

$$(2) \quad \hat{P}(\alpha) = \frac{\sum_{t=1}^T \hat{q}_t}{\sum_{t=1}^T \hat{q}_t} \alpha(t).$$

Using (1) and (2) it is easy to show that the concept of an insurance equilibrium is equivalent to that of a contingent claims competitive equilibrium. Under our assumption of risk aversion (which implies concavity of U) this also means an insurance equilibrium will exist and is a Pareto-optimal allocation of losses.

We should perhaps mention a few properties of insurance equilibria, and the corresponding distribution of losses, which are based on the individual equilibrium conditions:

$$(3) \quad \frac{u'_i(\hat{I}_i(t)) Pr^i(t)}{u'_i(\hat{I}_i(1)) Pr^i(1)} = \left(\frac{q_t}{q_1} \right) \quad \text{for } i = 1, \dots, n.$$

By combining (2) and (3), we find that

$$(3') \quad \begin{aligned} \hat{P}(\alpha) &= \frac{E^i[u'_i(\hat{I}_i(t))\alpha(t)]}{E^i[u'_i(\hat{I}_i(1))]} \\ &= \frac{COV^i[u'_i(\hat{I}_i(t)), \alpha(t)]}{E^i[u'_i(\hat{I}_i(t))]} \\ &\quad + E^i[\alpha(t)] \end{aligned}$$

where E^i and COV^i are computed using Pr^i .

If all individuals have the same probability distribution, i.e., if $Pr^i = Pr$ for all i , then we can drop the superscript on E^i and COV^i in 3'. In [5], Borch shows that if $Pr^i = Pr$ for all i , then $\hat{Z}_i(t)$ depends on $x_1(t), \dots, x_n(t)$ only through the total loss, $X(t) = \sum_{i=1}^n x_i(t)$, and that $\hat{Z}_i(t)$ rises with $X(t)$.

When i is risk averse this implies that $u'_i(\hat{I}_i(t))$ rises as $X(t)$ rises: i.e., $u'_i(\hat{I}_i(t))$ tends to be positively correlated with $X(t)$, the economy's total loss. This indicates that policies α which pay more as $X(t)$

risers will tend to cost more, because they will be likely to have a positive covariance with $u_i(I_i(t))$. This also should indicate why the price of a policy will depend on more than just its marginal probability distribution. In some of his work, especially [5], Borch seems to be searching for a price function that depends only on the marginal probability distribution of α , and it is not surprising that he does not find a satisfactory price function of this type. But his contentions that the failure to find such a price function means that a competitive price (of any sort) does not exist, or that game theory is necessary to describe results, seem to be invalid. There is, as we have shown, a notion of competitive equilibrium arising from a market, which is Pareto optimal. Whether *in empirical fact* insurances are priced in the way this model suggests is a separate question.²

II

Up to this point we have seen that it is possible to define an insurance equilibrium by essentially relabelling bundles of Arrow certificates. Insurance equilibrium is the same as an Arrow equilibrium and preserves the Pareto-optimal properties of such equilibria. At this point it is customary to comment on the large numbers of kinds of certificates needed for Arrow equilibrium and the obvious nonexistence (and probable impossibility) of such certificates

² To illustrate these results, suppose that all consumers are risk averse and exhibit constant risk aversion in the sense of Pratt [9] and Arrow [3]; i.e., $u_i(I_i(t)) = -e^{-a_i I_i(t)}$. If we let $A = [\sum_{j=1}^n a_j^{-1}]^{-1}$, and if $Pr^i = Pr$, for all i , then

$$\hat{z}_i(t) - \hat{z}_i(1) = a_i^{-1} A X(t).$$

It also follows that

$$\alpha_i(t) = \left[\sum_{j \neq i} a_j^{-1} \right] A x_i(t) - a_i^{-1} A \sum_{j \neq i} x_j(t)$$

and

$$\hat{P}(\alpha) = \{E[e^{AX(t)}]\}^{-1} E[\alpha(t) e^{AX(t)}]$$

Using this result it is easy to construct examples where $\hat{P}(\alpha) \neq \hat{P}(\alpha')$ even though α and α' have the same probability distribution.

in the world we observe. For example, Borch [6, p. 345] asserts that "it seems quite natural to assume that people have a fair idea of the equilibrium prices of ordinary commodities, and that they can figure out how much they want to buy or sell at these prices. It seems less natural to assume that people can figure out the equilibrium prices for 'Arrow-certificates.'"

In part, this conclusion seems to arise from a too literal interpretation of Arrow's model. People actually trade in insurance policies, not Arrow certificates, and it may not be unreasonable to assume that they have some notion of the going price for insurance policies of various types. However, the notion of an insurance equilibrium as we have defined it required that any policy α be available. This does not occur in actual insurance markets.

This may in fact be no difficulty if the class of policies available includes the optimal policies. We saw above that if all individuals share the same probability distribution over states that the optimal policies depend only on the total loss: i.e., the only policies that will be sold and observed are those in which the payment in two states will be the same when the economy's total loss is the same. That is, the number of states which need to be distinguished in optimal policies is equal to the number of different possible total losses ("social states"). The number of social states will ordinarily be less than T , since $\{1, \dots, T\}$ includes states which give the same total loss, but different distributions over individuals.

Note in particular that the size of the individual's initial own loss only affects the final loss he bears insofar as it affects the total loss. In the two-person, single-loss case analyzed by Arrow [2], an optimal arrangement is one in which an increment in the individual's loss will not be wholly compensated by the insurer, but in a many-person, many-loss context the size of the individual's loss will have al-

most no relationship with the loss he finally experiences, and the usual forms of "coinsurance" may not be optimal.

One may conceptualize the notion of optimal equilibrium as an exchange of Arrow certificates for all the states. But if the optimal outcome is known, this trading process can be, in a sense, "internalized" if some middleman simply offers optimal policies which need only distinguish social states. Individual exchanges do not have to be made if the optimal total policies can be marketed. As long as marketing or transactions costs are positive, we would expect this kind of internalization to occur.

It seems, however, that this is not really a complete explanation of real-world insurance outcomes for two reasons. First, there will still be a very large number of social states, and the shares of total losses that people may bear in social states may be different for different social states. Shares will not in general be the same for all social states. Second, many policies which are observed in practice do not even depend on the total loss, i.e., the typical situation is that a payment to i depends only on i 's loss, and so i 's final wealth position does not depend on the total loss. Thus it appears that the class of policies actually available are too restricted to allow for the achievement of Pareto-optimal allocations in the Arrow-Debreu models even if there is no common probability distribution, Pr .

In attempting to explain this aspect of insurance market behavior it seems reasonable for us to assume that more complicated policies would be more costly to provide. In particular we suppose that in order for the insurer, A , to provide insurance to person B , A has to know what B 's loss is; i.e., he has to have "information" about B . If there is a real cost to obtaining this information, the information may not be provided, in which case B

will not be insured by A . When there are a large number of people, say n , the cost of information may make it profitable for one person to obtain information about all of the others and insure them. However, the information cost may be sufficiently large so that it never pays for a person other than n to insure anyone. In such a case the policies held by $1, 2, \dots, n-1$ will depend only on their own loss.

By exploiting the framework provided by Radner in [10] we are able to introduce the lack of complete information into the analysis of insurance markets and to consider the costs of improving information. For example, consider a simple two-person, two-state insurance economy. Let $x_A(1) = x_A(2) = 0$ and $x_B(2) > x_B(1) = 0$, where $x_i(t)$ is person i 's loss in state t ; $i = A, B$; $t = 1, 2$. If person A only observes his own loss then he is unable to distinguish state 1 from state 2. If he attempts to provide the insurance policy α in return for $P(\alpha)$ he will have to rely on B to tell him whether state 1 or state 2 has occurred. If state 1 occurs it will be possible for B to claim that state 2 has occurred and collect $\alpha(2)$. As a result, an attempted trade from F to G will in fact move person A to point H in Figure 1. Unless A obtains the ability to distinguish $t=1$ from $t=2$ he will never make a trade.

Suppose that the economy's technology is such that the process of providing information consumes y units of income regardless of which state occurs.³ If y is not too large, i.e., if information is not too expensive, then it will be efficient for the economy to acquire information and have A insure B . Such a case is depicted in Figure 2, where the shaded area contains the Pareto-optimal points. On the other

³ This, of course, is a very special assumption. First, information may be costlier if one state occurs than another. Second, we assume that only perfect information is possible; i.e., that information is discrete.

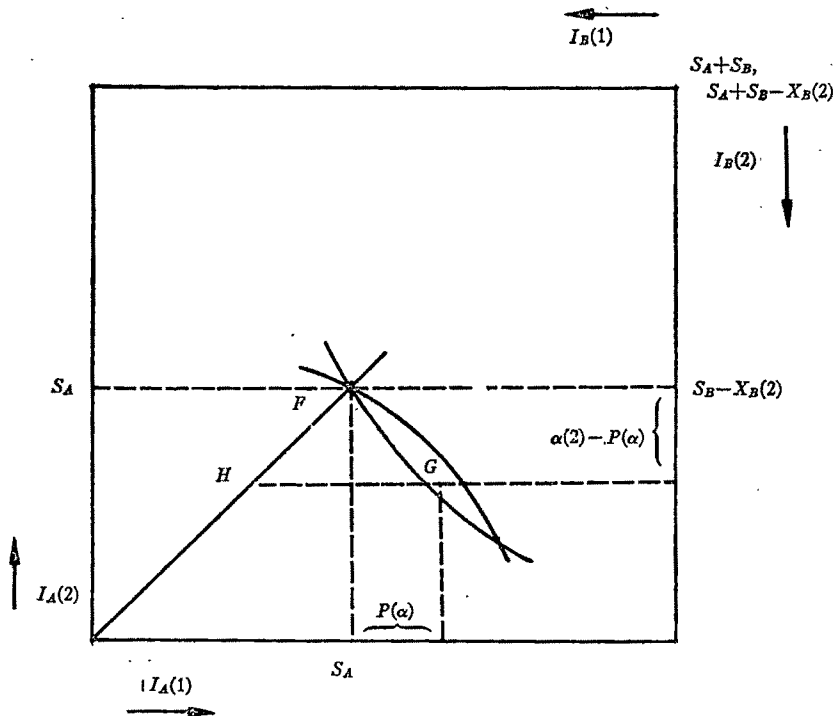


FIGURE 1

hand if information is expensive, if, for example, information costs y' , then no trade is Pareto optimal. Note that in such a situation it is possible for a Pareto-optimal allocation to be one in which there is no insurance, because the information cost of providing it is too high.

The cost of information is obviously a kind of transaction cost that must be paid if trade is to take place. It is one of the costs of having a market. When viewed in this way it should be clear that *A*'s acquisition of information about *B* is a public good (in the sense of a consumption externality) since both *A* and *B* benefit from the trade this acquisition made possible.⁴ This of course brings up the problem of the allocation of information, since the

usual competitive markets are known to allocate public goods in a nonoptimal way if they function at all.

A mechanism for obtaining optimal allocations in this example is to have A and B share the cost of information and then trade insurance policies competitively. One could imagine that along with the competitive price function $P(\alpha)$ (or the Arrow prices q_i), the market also announces the shares of y which A and B would bear. Both A and B would then maximize utility subject to their budget, while considering the shares and prices as parameters. Point K in Figure 2 is an equilibrium for this kind of process. Such an equilibrium would exist if, as the initial allocation moved from F (B bears the entire cost of information) to J (A bears the entire cost of information), the corresponding competitive equilibria enter the shaded region.

⁴ In analyzing a similar case, Radner [10] calls this cost a "set-up" cost. But it is really possible in many situations to consider "set-up" costs to involve public goods and vice versa.

combining (5) and (6) yields

$$(7) \quad \frac{u'_i(S_i - \hat{z}_i(x_i))}{u'_i(S_i - \hat{z}_i(x_i(1)))} = \frac{E[u'_n(S_n - \hat{z}_n(t)) \mid x_i]}{E[u'_n(S_n - \hat{z}_n(t)) \mid x_i(1)]}.$$

If we let $u_i = -e^{-a_i I_i(t)}$ and assume that x_i is independent of x_j then we can show that

$$\hat{z}_n(t) - \hat{z}_n(1) = x_n(t) + \sum_{j=1}^{n-1} a_j^{-1} A_j x_j(t)$$

and

$$\hat{z}_j(t) - \hat{z}_j(1) = a_j^{-1} A_j x_j(t),$$

where

$$A_j = [a_n^{-1} + a_j^{-1}]^{-1}.$$

IV

The number of individuals who actually become insurers and the form of actual insurance policies will, of course, depend of how the benefits from insuring in various ways compare with the information costs of the insurance and how those benefits and costs change as the number of insurers and insureds increase. Suppose, for example, that n individuals in a population of N identical persons agree to share

the social loss $\sum_{i=1}^N x_i = X^N$ equally.⁶ How

⁶ More generally, for any distribution of information let K_j be those individuals whose loss j observes and let \mathcal{K}_k be those individual's who observe k 's loss. If $u_i(I_i(t)) = -e^{-a_i I_i(t)}$ then j 's optimal loss share is

$$\hat{z}_j(t) = \sum_{k \in \mathcal{K}_j} \hat{z}_j^k(x_k(t))$$

where

$$\hat{z}_j^k(x_k(t)) - \hat{z}_j^k(x_k(1)) = \sum_{i \in \mathcal{K}_k} \left[\frac{a_j}{a_i} \right]^{-1} x_k(t).$$

These results depend crucially on the independence of the x_i 's, and the assumption that $P_T = P_T^i$, for all i .

⁶ Equal sharing may not be optimal, even in the limited sense of Radner, if persons are not identical.

does the benefit from having additional people insure the social loss change as n rises? We can use a result which Arrow and Lind [4] obtained in another context to answer this. Define $k_N(n)$ as the certain loss which has the same utility as the random loss $1/n X^N$. Arrow and Lind show that $n k_N(n) \rightarrow 0$ as $n \rightarrow \infty$, with N constant. This implies that the additional total value to the economy of having additional people share the social loss is virtually zero if n is large. So if the cost of giving information about the total loss to additional individuals is constant beyond some point, and "high" relative to X , the optimal number of persons who share the total loss will be less than the total population; i.e., n will be less than N . Some persons will be "insurers," some will not. For those who are not, insurers' final wealth will not vary with the total loss—a characteristic, as we noted earlier, of actually observed insurance arrangements.⁷ If, however, N is too small relative to X^N it may be true that for all $n \leq N$, $n k_N(n) - (n+1) k_N(n+1)$ —which equals the gain from providing an additional individual with information about the total loss—exceeds the cost of this information. In such a case all N individuals will be insurers. But will this be true as $N \rightarrow \infty$? Note that if we raise N we also tend to raise X^N so it is not immediately clear that by raising N we could reduce $(N-1) k_N(N-1) - N k_N(N)$ below the cost of providing information about the total loss to the last individual (suppose that this cost remains constant as N rises). If, however, the x_i 's are independent and identically distributed with mean μ and variance σ^2 we can show that $(N-1) k_N \cdot (N-1) - N k_N(N) \rightarrow 0$ as $N \rightarrow \infty$.⁸ So for

⁷ The person's final wealth may, for example, vary with his own loss, as under a coinsurance arrangement, or it may not vary at all, as under a full coverage arrangement.

⁸ If the x_i 's are independent and identically distributed with mean μ and variance σ^2 then

large N , one individual will not bear a share of the loss.

Generalizing this result leads to the conclusion that there are likely to be, say, m persons whose final wealth positions depend on the total loss, and $(N-m)$ persons who make a certain payment in return for full coverage of their own loss, so that their own incomes are certain. For these $(N-m)$ persons, their final wealth does not depend on the total loss.

Persons who bear part of the total loss might be thought of as having a kind of split personality in which they make a certain payment in return for coverage which does not depend on the total loss but in which they hold "stock" in an insurance "firm" which makes their final wealth positions vary with the total loss. This arrangement is probably preferable to the alternative in which, for each of the m persons who suffers a loss, the insurance payment received varies in such a way as to make his final wealth position depend

$$E \left[\frac{1}{N} \sum_{i=1}^N x_i(t) \right] = \mu$$

and

$$\text{var} \left[\frac{1}{N} \sum_{i=1}^N x_i(t) \right] = \frac{1}{N} \sigma^2.$$

In addition, suppose $u_i = u$ and $S_i = S$ for all i . If the third moment of x_i is finite we can use equation (5) in Pratt [9] to show that

$$k_N(N) = \frac{\sigma^2}{2N} r(S - \mu) + o\left(\frac{1}{N}\right)$$

where $r = -u''/u'$ and $No(1/N) \rightarrow 0$ as $N \rightarrow \infty$. Then

$$Nk_N(N) \rightarrow \frac{\sigma^2}{2} r(S - \mu)$$

as $N \rightarrow \infty$. Similarly for any fixed m

$$k_N(N-m) = \frac{\sigma^2 N}{2(N-m)^2} r\left(S - \frac{N}{N-m} \mu\right) + o\left(\frac{1}{N}\right)$$

so that

$$(N-m)k_N(N-m) \rightarrow \frac{\sigma^2}{2} r(S - \mu)$$

as $N \rightarrow \infty$.

on the total loss. For individuals as *prospective insurance buyers*, in this kind of Pareto optimum payments do not depend on the total loss. Insurance firms arise, in this context, as a way of economizing on information.

V

Before concluding we should mention the relationship between information costs and the problem of "moral hazard."⁹ In commenting on a paper by Radner [11], Arrow noted that the problem of moral hazard is in essence an information problem. Moral hazard arises when an individual has the ability to affect his loss in some or all states by taking some "discretionary" action. A "state" here describes all aspects of a situation outside the control of the individual, either in a certain or a probabilistic context.¹⁰ If complete information is available to the insurer the insurance payment can be a function of both the state and the discretionary activity, and optimal arrangements can be arrived at. In general, however, information about the loss is likely to be available at much lower cost than information about the state or the discretionary activity.

Conclusion

Consideration of insurance arrangements, in a world in which information is costly, provides us with some explanation of the existence of arrangements which, from the viewpoint of a theory for a world of costless information, would be nonoptimal. We have explored a few of the impli-

⁹ On the problem of moral hazard, see Pauly [8] and Arrow [2].

¹⁰ For example, a stranger dropping a lighted cigarette might describe a state, and a store-owner's expenditure on fire extinguishers the discretionary activity. Or an appendicitis attack might describe a state, and the use of a particular amount of medical care of a particular type the discretionary activity.

cations for optimal arrangements of introducing an information cost.

REFERENCES

1. K. J. Arrow, "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Rev. Econ. Stud.*, April 1964, 91-96.
2. ———, "Uncertainty and the Welfare Economics of Medical Care," *Am. Econ. Rev.*, December 1963, 941-74.
3. ———, *Aspects of the Theory of Risk-Bearing*, Helsinki, 1964.
4. ———, and R. C. Lind, "Uncertainty and the Evaluation of Public Investment Decisions," *Am. Econ. Rev.*, June 1970, 364-78.
5. Karl Borch, "Equilibrium in a Reinsurance Market," *Econometrica*, July 1962, 424-44.
6. ———, "Economic Equilibrium under Uncertainty," *Int. Econ. Rev.*, October 1968, 339-46.
7. G. Debreu, *Theory of Value*, New York, 1959.
8. M. V. Pauly, "The Economics of Moral Hazard," *Am. Econ. Rev.*, June 1968, 531-36.
9. J. W. Pratt, "Risk Aversion in the Small and in the Large," *Econometrica*, January-April 1964, 122-36.
10. R. Radner, "Competitive Equilibrium under Uncertainty," *Econometrica*, Jan. 1968, 31-58.
11. ———, "Problems in the Theory of Markets under Uncertainty" (and discussion by Mack, Arrow, and Rader), *Am. Econ. Rev.*, May 1970, 454-64.

Insurance, Information, and Individual Action

By MICHAEL SPENCE and RICHARD ZECKHAUSER
Harvard University

This paper looks at some intricacies and difficulties that arise in the real world operation of contingent claims markets. Insurance contracts, the most readily observable and perhaps most important example of contingent claims markets in action, provide the focus for our discussion.

The purpose of insurance is to protect risk-averse individuals from suffering the full consequences of those actions on the part of nature which affect them unfavorably. The parties to an insurance contract agree that when the actions of nature become known, those most favorably affected will transfer resources to those who turn out to be less fortunate. If the contract is to provide protection in this way, it is essential that there be (at least substantial) independence in the actions nature takes with respect to different insured individuals. Such independence is assumed for the remainder of this paper. This enables us to look at insurance schemes from the standpoint of a single individual who is representative of the many who are insured under a contract. A further assumption is implicit in this approach—it is that all individuals have identical prospects, resources, and utility functions.¹

¹ This assumption is not restrictive in the manner it might appear. If individuals do differ with regard to any of these variables, the situation can be thought of as one in which nature has already taken a move. In future efforts we hope to get into problems in which there is more than a single exchange of moves between nature and the individual. We also intend to expand the range of actions available to the individual. For example, if nature makes moves before the insurance contract is drawn up, some individuals may choose not to participate. This is the problem of adverse selection. It is not considered here.

The possible arguments of our representative individual's utility function, u , are his overall wealth level, w , the act of nature, n , and his individual action, a . His utility function has the familiar von Neumann-Morgenstern properties; it enables him to make choices over lotteries on all the arguments of his utility function. The individual is assumed to be risk averse with respect to lotteries on wealth.

We classify insurance schemes according to their structural characteristics in three areas. The first is (1) the presence or absence of individual choice. There is individual choice if a is a nontrivial argument in the utility function. The individual choice model is quite general; thus, the choice variable, a , might represent a level of investment, including investment in self, or the purchase of some specific good, say medical services.

Where there is room for individual choice, we will be concerned with (2) the sequencing of moves between the individual and nature, and (3) the information state, S , monitored by the insurer. This information state, which may be a vector, will be a function of the act of nature and the action of the individual.

The insurance scheme works by having the insurer determine a monetary payoff that he makes to the individual, the size of the payoff to depend on the information state he monitors. The insurance scheme is fully described by what we call the insurance payoff function, $g(S)$. It is subject to the actuarial constraint that it have a break-even financial expectation. The insurer's object is to maximize the repre-

sentative individual's expected utility subject to this constraint.

Our principal concern in this paper is to examine how (2) and (3) interact in the determination of optimal insurance schemes. As a standard of comparison, we consider first simple insurance—insurance situations in which there is no room for individual choice.

*Case I: No individual Choice—
Insurer Monitors n*

If individual choice plays no role, the sole determinant of the individual's utility is nature's action, n . The distribution of n is given by nature's density function, $f(n)$. We treat n as a continuous variable in this paper, but with minimal modification our results hold equally well if n is discrete.

The insurer monitors the action of nature; in this case $S = [n]$. He gives the individual a payoff $g(n)$. This payoff added to the insured's initial wealth, w_0 , gives the wealth argument of his utility function. The individual's expected utility under this scheme is

$$(1) \quad \int u(w_0 + g(n), n) f(n) dn.$$

The break-even constraint for this scheme is

$$(2) \quad \int g(n) f(n) dn = 0.$$

The insurer's objective is to maximize (1) subject to (2). He can employ the calculus of variations to derive the marginal efficiency condition for the optimal insurance payoff function. That condition is that there be a constant λ such that

$$(3) \quad u = \lambda$$

or $f(n) = 0$, for all values of n . The optimal $g(n)$ keeps the marginal utility of income constant. Its dependence on $f(n)$ is only through the value of the parameter λ ,

which is determined by the constraint equation (2). Thus, the shape of $g(n)$ is only indirectly affected by nature's density function. These properties of the optimal insurance payoff function are characteristic of all cases where the insurer can monitor all available information.

A special subcase of interest is one where the act of nature itself is a monetary payoff so that $u(w, n)$ can be written $v(w + n)$. The optimizing condition becomes

$$(4) \quad v'(w_0 + n + g(n)) = \lambda.$$

This implies that $g(n) = k - n$, where k is an arbitrary constant. The integral constraint, (2), requires that $k = \bar{n}$, the mean of n . The optimal insurance scheme when an individual has an uncertain income over which he has no control is one which always gives him his expected income.

We turn, for the remainder of this paper, to cases where the insured individual, in full knowledge of the insurance payoff function, takes an optimizing action, a .

*Case II: Individual Chooses Before Nature
—Insurer Monitors R and a*

For simplicity, we consider a model in which the insured's utility function has but one argument, *ex post* wealth. This model requires that the individual's action be convertible to a monetary equivalent. It seems most reasonable to think of a as an investment the individual makes to increase an uncertain monetary return he will receive. We represent this return as $R = r(n, a)$.² The insurer monitors this monetary return as well as the individual's

² Alternatively, we may think of a as a parameter which alters the distribution of the payoff R . Since $R = r(n, a)$ and n has a distribution, $f(n)$, we may write n as a function of R and a ,

$$n = h(R, a).$$

The conditional distribution of R given a is

$$t(R|a) = f(h(R, a)) | h_R(R, a) |.$$

action in determining his payoff; that is, $S = [R, a]$. The individual's wealth after he makes his investment, receives his monetary return, and obtains his payoff from the insurer is

$$(5) \quad w = w_0 + R + g(R, a) - a.$$

The insured will undertake the action that maximizes his expected utility. That is, given the insurance scheme and $f(n)$, he will pick a to maximize his expected utility

$$(6) \quad \int u(w_0 + r(n, a) + g(r(n, a), a) - a)f(n)dn.$$

Here R is written in its functional form, $r(n, a)$, to emphasize its dependence on a .

An individual who invests in his earnings future by obtaining an education, or a person who fireproofs his home to make it less likely to be damaged or destroyed, is an individual taking an action which together with the act of nature determines the monetary return he receives. The individual has some control over his fate in these circumstances, but the additional influence of the uncontrollable action of nature means he cannot determine it completely. Thus, it may be desirable to institute an insurance plan.

Let us look at the problem of the insurer who wishes to maximize the insured's expected utility. The break-even constraint is that

$$(7) \quad \int g(r(n, a), a)f(n)dn = 0$$

for the particular a that maximizes (6) for the given $g(R, a)$. (Note that this constraint need not be satisfied for other values of a , because they will never be chosen and hence are irrelevant to real-life actuarial considerations.)

The insurer wishes to pick $g(R, a)$ to maximize (6) subject to this constraint.

His problem is simplified and his performance (in terms of achieved expected utility) improved because he can monitor a and employ it as an argument of his payoff function. He can make the insured select $a = a^*$ by making $g(R, a)$ sufficiently negative for $a \neq a^*$. Given this control over the insured's action, the insurer's problem becomes to find the pair a^* together with $g(R, a^*)$ that maximizes (6) subject to (7). The optimal $g(R, a^*)$ would be determined, as before in the no-individual-choice case, by the condition

$$(8) \quad u' = \lambda,$$

which implies that

$$(9) \quad g_1(R, a^*) = -1.$$

The important point to realize, for this case, is that the adverse incentives problem is eliminated because the insurer can monitor the insured's action and structure the insurance payoff function so that the selected a will equal a^* .³

³ If the insurer had the ability to monitor only n , the outcome that would be achieved would be identical. The insurer would select the optimal insurance payoff function, call it $g^1(n)$, for which there would be a corresponding optimal act for the insured, call it a^* . Expected utility is U^1 .

Assume that g^1 is unique and suppose that the insurer could improve expected utility by monitoring $r(n, a)$ and a . Represent his optimal payoff function-action pair as $g^2(r(n, a), a)$ and a^{**} , with expected utility U^2 . Now let $g^3(n) = g^2(r(n, a^{**}), a^{**})$. This new insurance payoff function satisfies the break-even constraint; it gives the same distribution of payoffs as did g^2 . Let the insured select his optimal act for g^3 , call it a^{***} . The expected utility for this pair $U^3 \geq U^2$. But g^3 is an admissible insurance payoff function which monitors only n . Because g^1 was optimal, $U^1 \geq U^3$. The previous weak inequalities imply that these expected utilities are equal. A parallel argument shows that expected utility cannot be greater when $S = [n]$ than when $S = [r(n, a), a]$. Thus we have

$$(a) \quad g^1(n) \equiv g^2(r(n, a^{**}), a^{**}) \equiv g^3(n),$$

and

$$(b) \quad a^* = a^{**} = a^{***}.$$

If the insurer can monitor n only, he will

$$(c) \quad \text{maximize}_{g(n)} \int u(w_0 + r(n, a) + g(n) - a)f(n)dn$$

*Case III: Individual Chooses Before Nature
—Insurer Monitors Only R*

Frequently it is impossible or prohibitively costly for the insurer to monitor a as well as R . The term a may represent, for example, the effort an individual makes to improve his earnings opportunities or the degree of care with which he drives his car. When a cannot be observed directly, the information state contains only R , $S = [R]$, and the insurer's payoff function takes the more restricted form $g(R)$. The opportunity provided before to enforce a choice of a is no longer available.

The insurer's optimization problem becomes more complex. He selects $g(R)$ to maximize

$$(10) \quad \int u(w_0 + r(n, a) + g(r(n, a)) - a)f(n)dn,$$

subject to two constraints. The first is the usual break-even condition

subject to

$$(d) \quad \int g(n)f(n)dn = 0$$

and

$$(e) \quad \int (r_2 - 1)u'(n)dn = 0,$$

where the second constraint represents the maximizing action of the individual. We know that u' is constant for the optimal scheme, so that the individual's maximizing a satisfies the equation

$$(f) \quad \int (r_2 - 1)f(n)dn = 0.$$

The expected net return from an additional dollar of investment is zero.

Not only does the insurance scheme spread risk, it also makes the investment decision productively efficient by overcoming the distortionary impact of the individual's risk aversion. In general, when there is no insurance, the marginal condition (f) will not be satisfied. In cases where a is a defensive investment against low values of w , as it is say with safety measures, the selected a will be too large. On the other hand, if a generates high payoffs with low probabilities, the chosen a will be too small. Perhaps investment in education for disadvantaged individuals represents the latter situation.

$$(11) \quad \int g(r(n, a))f(n)dn = 0.$$

The second is that a is selected to maximize (10) given $g(R)$. This second constraint can be given by the marginal condition⁴

$$(12) \quad \int [r_2 + g'r_2 - 1]u'f(n)dn = 0.$$

Treating the problem again as one in the calculus of variations, the marginal condition for the optimal $g(R)$ function is

$$(13) \quad \lambda u'' + u' \left[1 + \lambda \frac{d}{dn} \left(\frac{r_2}{r_1} \right) + \lambda \frac{r_2}{r_1} \frac{f'}{f} \right] - \Phi = 0,$$

where λ and Φ are parameters, determined by the constraints (12) and (11).

In this case, g depends directly on the distribution $f(n)$, not just through the parameters λ and Φ . There is a significant difference from the previous cases. It is no longer possible to keep the marginal utility of wealth constant. The $g(R)$ function must be such that on the whole *ex post* wealth is an increasing function of the payoff received. Otherwise the insured individual would have no incentive to undertake any action at positive cost, no matter how favorably that action would affect this payoff. To achieve appropriate incentives, the insurance plan must sacrifice some of its risk-spreading capabilities.⁵

⁴ This will be true, for example, if $u(w_0 + r(n, a) + g(r(n, a)) - a)$ is everywhere concave in a .

⁵ A simple example will make clear the loss of efficiency due to the insurer's inability to monitor a as well as R . The insured's utility function is $\log(w)$, with initial wealth given. The monetary payoff he receives, is functionally defined as $R = (n \cdot a)^{1/2}$. Nature's action is determined by a density function which is uniform on some interval. The form of the optimal payoff function is

$$g(R) = \alpha - R + \beta(1 + \gamma R)^{1/2},$$

where $\alpha, \beta > 0$, and $\gamma > 0$ are parameters whose values

*Case IV: Individual Chooses After Nature
—Insurer Monitors n and a*

We turn our attention now to situations in which the insured individual takes his maximizing action, a , after he learns nature's move. The individual who knows his wage rate (as determined in a lottery conducted by nature) and decides how much he wishes to earn, or the individual who knows his medical condition and decides how much to spend on medical care would be a man in such a situation.

The insurer's monitoring capability enables him to employ a payoff function $g(n, a)$. Insurance mechanisms that fall into this category may be thought of as unfamiliar versions of the common purchase option. Medical insurance reimbursement schemes, for example, state how much an individual must pay if he purchases a given amount of medical service, depending on his medical condition. An income tax scheme endowed with this enriched information-monitoring capability would be empowered to make an individual's tax assessment depend on his earning capability as well as his level of earnings. The individual's action in this case is to sacrifice income to purchase leisure. The tax scheme would tell him how much a given leisure purchase would cost him (income given up less taxes saved) as a function of his wage rate.

We will demonstrate below that if n can be known, the additional ability to monitor a is of no value. This means that there is no loss of efficiency if the payoff from medical insurance is made solely a function of the insured's condition, or if the imposition of an income tax is made to depend only on an individual's probabilistically-determined wage rate.

can be found by direct search.

A loss of efficiency comes about because dg/dR is not -1 , complete risk spreading is not achieved, as it is when the insurance payoff function can monitor both a and R .

In this case, the actions of both nature and the individual enter the insured's utility function, $u(w, n, a)$. We assume that u is strictly concave in its first and third arguments. After n is announced, the individual selects a to maximize $u(w_0 + g(n, a) - a, n, a)$. Given n and $g(n, a)$, the insured's efficiency condition for the choice of a is

$$(14) \quad u_3 + u_1(g_2 - 1) = 0.$$

The insurer's optimizing problem is to

$$(15) \quad \underset{g(n,a)}{\text{maximize}} \int u(w_0 + g(n, a) - a, n, a) f(n) dn,$$

subject to the condition given in (14) and the break-even constraint

$$(16) \quad \int g(n, a) f(n) dn = 0.$$

Because the insurer can monitor n , he can force the insured to select whatever combination of a and w he wants. The system is in this sense controllable. Suppose that the insurer wants $a = p(n)$ and $w = w_0 + g(n, a) - a = q(n)$. He may then set

$$(17) \quad g(n, a) = \left[1 - \frac{u_3(q(n), n, p(n))}{u_1(q(n), n, p(n))} \right] \cdot [a - p(n)] + q(n) + p(n) - w_0.$$

Since

$$(18) \quad g_2 = 1 - \frac{u_3(q(n), n, p(n))}{u_1(q(n), n, p(n))},$$

and u is strictly concave in a and w , the insured will select $a = p(n)$ because of condition (14) above. When he does, his wealth will be $q(n)$ as required.

This controllability, which is due to the ability of the insurer to monitor n , simplifies the problem considerably. For the

insurer may now substitute the following problem:

$$(19) \quad \text{maximize} \int_{p(n), q(n)} u(q(n), n, p(n)) f(n) dn$$

subject to

$$(20) \quad \int [p(n) + q(n) - w_0] f(n) dn = 0.$$

Employing the standard calculus of variations formulae, he can find the two marginal conditions

$$(21) \quad u_1 = \lambda,$$

and

$$(22) \quad u_3 = \lambda.$$

They imply that the marginal utility of income and dollars expended on services are equal and constant at the optimum. More importantly,

$$(23) \quad g_2 = 1 - \frac{u_3}{u_1} \equiv 0,$$

so that

$$(24) \quad g(n, a) \equiv m(n),$$

for some function $m(n)$. This proves our earlier contention that if the insurer can monitor the act of nature directly, there is no additional gain in being able to monitor the action of the individual. In other words, the achievable expected utility is the same when $S = [n]$ as when $S = [n, a]$.

*Case V: Individual Chooses After Nature
—Insurer Monitors Only a*

We conclude with an examination of the case where the insurer can monitor only a , that is where $S = [a]$. An income tax scheme where the taxing authority can monitor income, but not earning opportunities, or a medical insurance plan that relates only to amounts spent, but not medical condition, would represent such a situation.

Here a is a signal for n . With the appropriate monotonicity properties, it may even be a perfect signal. But due to the problem of adverse incentives, an insurance payoff scheme that monitors a as a signal for n will perform less well than one which can monitor n directly.

This problem is familiar in the context of purchase-option schemes. Income tax plans distort incentives for work; medical insurance reimbursement programs lead individuals to overexpend on medical services. This problem also has interesting ramifications in cases where a is not most appropriately interpreted as a dollar expenditure.

Assume that the amount of sleep an individual chooses is a perfect signal for his medical condition, the more sleep the more serious the condition. If an insurer could not monitor medical condition directly, nor anything else but the individual's chosen quantity of sleep, he would be forced to employ sleep as the sole argument of his insurance payoff function. This would give the insured an incentive to select a supraoptimal amount of sleep no matter what his condition. (We are assuming the insurer would like to give a higher monetary payoff the more serious the condition.) The sleep amount chosen might even remain a perfect signal for medical condition, but there would be no way to overcome the problem that differential payoffs for different amounts of sleep would introduce a new and inappropriate factor into the sleep-quantity decision.

The insurer in Case V recognizes the problems that confront him. He proceeds to optimize in a second-best world, one where the information he would like to monitor is not available. He will choose his insurance payoff function $g(a)$ to achieve the optimal tradeoff between the conflicting goals of furthering risk spreading and providing appropriate incentives. To define this function he can employ the Hamil-

tonian method with U , achieved utility, serving as the state variable, and g serving as the control variable.⁶ For fixed n , a is defined implicitly as a function of U and g by the equation

$$(25) \quad U = u(w_0 + g(a) - a, n, a).$$

The optimizing marginal condition for the insured's choice of a is

$$(26) \quad u_3 + u_1(g' - 1) = 0.$$

This may be written in the equivalent form

$$(27) \quad \frac{dU}{dn} = [u_3 + u_1(g' - 1)] \frac{da}{dn} + u_2 = u_2.$$

We let $\Phi(n)f(n)$ be the multiplier corresponding to the state variable U , and λ be the multiplier corresponding to the break-even constraint

$$(28) \quad \int g(a)f(n)dn = 0.$$

With this notation, the Hamiltonian becomes

$$(29) \quad H = [U - \lambda g + \Phi u_2]f.$$

The optimizing equations are therefore

$$(30) \quad \lambda g' = \Phi[u_{23} + u_{12}(g' - 1)],$$

$$(31) \quad \frac{d\Phi}{dn} = -\frac{f'}{f}\Phi - 1,$$

and

$$(32) \quad u_3 + u_1(g' - 1) = 0.$$

We can interpret these results in the concrete context of the medical insurance

⁶ A mathematically similar problem in the context of income taxation is treated more completely and rigorously by James A. Mirrlees, "An Exploration in the Theory of Optimal Income Taxation," to appear in *Econometrica*. He uses variational methods in large part, but indicates that the Hamiltonian method can be used in this way. The problems are structurally different in that n , the act of nature, does not appear in the control function $g(a)$ in our problem, whereas it does in his income tax problem.

example. Let n represent medical condition, with larger values representing greater illness. To put the problem in a more tractable form, so as to provide us with further insight into the properties of the optimal payoff function, we consider the case where wealth enters the utility function in an additive way. For this case, $u_{12} = u_{13} = 0$. Thus, condition (30) becomes

$$(33) \quad g' = \frac{\Phi}{\lambda} u_{23}.$$

Typically, u_{23} will be positive, the marginal utility of medical expenditure will increase with an increase in the severity of illness. This implies that g' will be positive as well, in contrast to the situation where the insurer was able to monitor n directly and $\partial g/\partial a = 0$. When only a can be monitored, the optimal insurance plan will not require the insured to pay the full marginal cost of his purchases; he will purchase too much.⁷

From equation (31), we note that $\Phi(n)$ depends entirely on $f(n)$, the distribution of the act of nature. We wish now to get some feeling for the shape of the optimal function $g(a)$. Differentiating (33) with respect to n and solving for the g'' gives

$$(34) \quad g'' = \frac{\Phi'}{\lambda} \frac{u_{23}}{a} + \frac{\Phi}{\lambda} u_{223} + \frac{\Phi}{\lambda} \frac{u_{233}}{a}.$$

The second order conditions for the insured's optimizing decision imply that $a > 0$. Moreover, it seems reasonable to assume that $u_{233} > 0$, meaning that the rate at which diminishing returns to medical services set in decreases with ill health. Similarly $u_{223} > 0$ is reasonable.

The sign of g'' therefore depends upon both the sign and magnitude of Φ' which

⁷ If the insurer can monitor only the action of the individual and if there is to be any risk spreading at all, the insured must be reimbursed somewhat as his level of expenditure increases. He will not be paying the full marginal cost of his purchases.

in turn depends on the distribution $f(n)$. If $\Phi' > 0$, then $g'' > 0$ and g is convex at that point. However, if Φ' is sufficiently negative, then $g'' < 0$ and g is concave at the relevant point.

We can posit one situation in which we would expect $g(a)$ to be concave, at least over a limited range of values. From equation (31), Φ' will be very large and negative if f'/f is very large and positive. In those regions of a corresponding to the regions of n where f'/f is very great, there will be a sharp increase in the frequency with which a is chosen as its value gets larger. If this frequency increase is sufficiently great, the relative importance of providing appropriate incentives as opposed to furthering risk spreading may increase as a itself increases. This will require that $g(a)$ flatten out somewhat and thus be concave over this range of values for a .⁸

For the most part, however, we would expect $g(a)$ to be convex.⁹ The risk spreading objective will receive relatively greater emphasis when expenditures are already high, for the greater is the individual's a , the less is his wealth, and the greater the utility cost to him of further expenditure.

⁸ By squishing together the upper values of the n scale, it will always be possible to have f'/f increasing rapidly. However, this will have the effect of increasing the rate of change of a with respect to n , and the two changes will cancel out, having no net effect on the shape of the optimal $g(a)$.

⁹ We note that if $g(a)$ is convex over most of its domain, then deductible policies, which are a common form of insurance for automobiles and health, may be a reasonable approximation to the optimum.

Conclusion

Some general principles emerge from our brief examination of different cases of insurance. If the insurer can monitor n directly, as he can with Cases I and IV, the insurance scheme can operate like a traditional contingent claims market. Full risk spreading can be achieved, and there will be no need to worry about adverse incentives.

Similarly, if the insurer can monitor the individual's action taken in advance of nature's act (Case II), the adverse incentives problem can be avoided by structuring the insurance payoff function to enforce the choice of the appropriate a .

These hopeful results do not hold, unfortunately, for Cases III and V. In these cases a signal which depends in part or completely on the insured individual's action is employed as the sole argument of the insurance payoff function. The insured will be induced to alter his natural maximizing action somewhat in order to influence this signal and thus increase his payoff from the insurer. The insurer can be cognizant of this adverse incentives problem, but he cannot overcome it. Given his limited information-monitoring capability, his selection of the optimal insurance payoff function is a second-best exercise. Neither complete risk spreading nor appropriate incentives for individual action will be achieved. To find the optimal mixture of these two competing objectives is a difficult problem, here as in the real world.

DISCUSSION

MARTIN J. BAILEY: The Spence-Zeckhauser paper presents a theory of optimum insurance provisions when all premia are actuarial: utility-maximizing zero-cost insurance, with and without moral hazards (which imply constrained, second-best optima), variously specified. Its assumptions are clear, its analysis thorough and apparently rigorous. My complaints are almost editorial in character. First, the insurer's utility function is implicitly risk-neutral; in his oral presentation, Spence clarified this point partly by saying that the insurer is the government. Thus the analysis covers only the demand side of the picture, and that only for the case of actuarial premia.

Further, to generalize this material the authors should consider the supply side of insurance and the response of demand to loading charges. They might wish to consider private insurers. Moreover, even on the government side, the assumption of no risk aversion is controversial, to say the least, and deserves discussion.

Kihlstrom and Pauly start out with the potentially fruitful idea that the insurance problem is logically the same as the problem of portfolio choice and of Arrow's market for contingent claims. Rigor and thoroughness are regrettably much less apparent in their paper than in the preceding one, however, and the second half has little connection with the first. They leap from the alleged failure of Borch's model to predict actual insurance market behavior to an explanation in terms of information costs, without reviewing other possible explanations, without serious discussion of the evidence, and indeed without describing Borch's assumptions and analysis. Although information cost deserves the kind of attention they give it, using it to "explain" every apparent discrepancy between theory and experience is a cop-out similar to giving the name "technological change" to every discrepancy between actual and explained economic growth. While seeming to explain everything, it explains nothing, when it offers no testable hypothesis. (The discrepancy that concerns them is that Borch predicts that each claim-

ant's policy would make his reimbursement of loss dependent on total insured losses, which on the face of it is not the way that insurance policies are written. As an afterthought, however, I would point out that insurance companies adjust premia by experience and set their premia to provide reserves for years of exceptionally high claims. Moreover, if a company's losses are large enough, it fails, and the claimants get less than their policies promised. Through these adjustments each policyholder shares in total insured losses.)

Although the Kihlstrom-Pauly discussion of information cost has intuitive appeal, the idea of applying the well-developed theory of risk to the particular case of insurance has by far the most unexploited potential. In their paper Brainard and Dolbear seem to recognize this point and present a good review of the broad sweep of the subject. However, to follow them one must get on his horse and ride off in all directions; their paper tries to do too much.

In one section Brainard and Dolbear chide Arrow for his superficial argument using a riskless discount rate when evaluating proposed government investments. Although I share their view, I would point out that Arrow's work falls into two distinct classes: first, there is a massive collection of brilliantly unassailable work that has enriched our understanding of many problems; second, there are some ideas that he keeps repeating, and that we can expect that he will go on repeating (although we can't predict what reasoning he will use to support them) regardless of the brilliantly unassailable counter-arguments that the rest of us bring against them. It is as much a waste of time to argue with him about the second group of ideas as it is about the first. Moreover, after coming up with one of these brilliantly unassailable counterarguments, we find to our chagrin that others have already come up with it at least twice before.

The rest of the Brainard-Dolbear paper offers a commendably comprehensive view of the insurance market, among other things, properly noting its similarity to markets for other risky claims (as did Kihlstrom-Pauly).

One wishes that someone would join these ideas with the power and elegance shown by Spence and Zeckhauser. A single paper containing the best features of all three papers would be a very good paper. As matters stand, the subject is a ripe apple, waiting to be picked.

HYMAN P. MINSKY:

"By 'uncertain' knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory Bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth-owners in the social system in 1970. About these matters there is no scientific basis on which to form any capable probability whatever. We simply do not know. Nevertheless, the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good Benthamite calculation of a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed."—J. M. Keynes, *Quart. J. Econ.*, Feb. 1937.

The above is how Keynes in his rebuttal to Viner's review characterized the uncertainty relevant to the behavior of a capitalist economy. It followed from this view that wealth was "a peculiarly unsuitable subject for the methods of the classical economic theory."

The papers under consideration take the Arrow-Debreu state preference approach to uncertainty seriously and generally speaking apply the model to determine whether insurance contracts can be designed so as to realize classical optimal situations. I will examine whether the Arrow-Debreu model can throw light on how uncertainty enters into macro-

economic theory and analysis, in particular the generation of social risk.

Even though this session is on social risk, only Brainard-Dolbear define and discuss the concept. They define social risk as "uncertainty in the aggregate consumption opportunities of society." As Kihlstrom-Pauly point out, in an early version of their paper, in the standard Arrow-Debreu scheme there is "an objective probability distribution $P_r(\cdot)$ over the set $\{1 \dots T\}$ of states and that this is known to all consumers." Thus the Brainard-Dolbear definition merely extends the set of states of nature and associated probabilities to include variable aggregate consumption levels.

A view of what social risk is follows from the Keynes citation. Even though at each instance decisions are made as if Benthamite calculations are appropriate, in truth each actor knows that for many problems his views are based on a limited capacity to scan alternatives and vague conjectures as to the likelihood of the various alternatives taking place. Thus, a degree of confidence is associated with the assigned probabilities, and actors recognize that states of nature, pay-offs and probability distributions other than those established are possible. Events will modify views as to possible alternatives and relevant probabilities, and these events need not be markedly different from events which would not have had this effect. Thus the behavioral relations that depend upon these changeable subjective conjectures are unstable.

This need to make decisions in the face of known instabilities leads to social risk. Brainard-Dolbear seem to recognize this in their asides on cyclical phenomena. The other authors are not directly concerned with these issues, although the cost of information considered by Kihlstrom-Pauly suggests why subjective evaluations cannot be avoided.

This is 1970. The financial happenings of 1966 and 1970 (the crunch and the squeeze) have instructed us in how financial events induce changes in the contemplated states of nature and the probability associated with each state. However, these events did not happen in a vacuum. The ground was set by strong expansions that were largely financed by portfolio adjustments. In both instances the Federal

Reserve induced rundowns in particular bank liabilities and forced banks and nonbanks into novel financial instruments for position making. Increasingly, liability structures were generated whose viability depended upon the fulfillment of buoyant income expectation and the proper functioning of special financial markets. In both years, fears arose that market failure was not only possible but imminent. Even though such failure was aborted by prompt Federal Reserve action, a revision of the states contemplated and their associated probabilities occurred.

Both 1966 and 1970 were preceded by periods of rapidly increasing investment demand. Liability experimentation more than offset the modest attempts by the authorities to restrict credit. Given the economy's history of business cycles, the initial set of likely states of nature and probability distribution induced conservative portfolio choices. The run of success modified both the alternatives considered and the associated probability distribution. Prolonged success of the economy erodes the felt uncertainty and thus both induces an investment boom and, by shifting liquidity preference, helps to generate the needed financing. Central bank action is ineffective in constraining such expansions unless it can increase felt uncertainty. In such circumstances, policy needs to bring the economy to the brink of failure of some financial markets.

Once successful brinkmanship increases the felt uncertainty, it is unlikely that an easing of financial dangers will lead to a quick resumption of strongly expanding investment demand.

The papers show that Arrow-Debreu format can yield insights into the design of insurance schemes, and may help make precise the obstacles to the achievement of optimality by way of insurance. However, for many phenomena the ingredients to a state-preference view are both subjective and unstable. Nevertheless, the Arrow-Debreu schema is valuable, for it allows us to inquire into how the subjective states of nature and their probabilities are determined. In particular, it enables us to state the Keynesian ideas about uncertainty more precisely. The introduction of uncertainty indicates that the scale of investment can

"fluctuate for reasons quite distinct . . . from those physical conditions of technical capacity to aid production which have usually been supposed hitherto to be the chief influence governing the marginal efficiency of capital" (Keynes, *op. cit.*) and thus it is a necessary ingredient to business cycle theory.

DONALD D. HESTER: One of the risks which a discussant at these meetings incurs is that papers will not be available before a session. The only completed paper in my hands is that of Messrs. Kihlstrom and Pauly. The audience will have to determine whether the state of the other two papers makes this discussion a totally uninsurable loss or a Pareto-efficient solution given the apparent high costs of distributing information. I shall, nevertheless, discuss all three papers, no matter how preliminary, in order to insure that Kihlstrom and Pauly will not be victims of moral hazard.

The paper by Spence and Zeckhauser differs from the other two papers on several counts. In particular it does not attempt to exploit closely the Arrow-Debreu framework. Using an information structure approach, common in the theory of games, the authors determine under what conditions benevolent insurance companies will be able to maximize the expected utility of insured individuals. It is not a team problem, however, for insured individuals may not always gain by disclosing information about their actions or the state of nature to an insurance company. Not surprisingly, they report that insurance companies will not maximize the expected utility of policy holders when information is incomplete or untimely.

The authors' approach is very restrictive, for they are forced to assume that all individuals have identical prospects, resources, and utility functions. The definition of the utility function itself varies throughout the paper, and its arguments, the state of the world, an individual's action, and his overall wealth often do not have an easily interpretable metric.

Spence and Zeckhauser err when they say that it is essential to assume that the actions of nature against different insured individuals must be independent for contracts to provide

protection. As a counter example, an insurance company could accommodate the players of two-cylinder Russian roulette, a game whose outcomes are perfectly negatively correlated.

Finally, they assume away a problem which is at least indirectly covered in the remaining papers of this session, the feasibility of offering complete insurance. Spence and Zeckhauser constrain their insurance company to break even actuarially; that is, they require that the expected value of insurance company profits be zero. This does not mean that an insurance company will in fact break even, and this possibility needs to be introduced into their analysis.

By way of contrast, the paper by Brainard and Dolbear is not so much concerned with the information structure of the insurance problem as it is with the question of whether or not enough different insurance contracts will be offered to reach Pareto optimality in markets for contingent claims. They make useful contributions (1) by carefully distinguishing between social and private risk and (2) by translating the Sharpe model of security risk into the framework of a contingent commodity world. They then conjecture with admittedly sparse evidence that the options available in asset markets are small in number and probably not sufficient to span the space of contingent commodities. They conclude by examining human capital and suggest that contingent claims in these markets are unlikely to develop because of the presence of moral hazard.

I am inclined to accept their verdict that the set of available assets is insufficient to span the space of contingent claims. I do not, however, regard this as evidence of Pareto inefficiency in any meaningful sense, for they, like Spence and Zeckhauser, have not considered the costs of sustaining asset markets. With a sufficiently effective surveillance sys-

tem, the problems of moral hazard can be overcome, but such information systems are costly and surely endanger privacy. If the cost of spanning the space of contingent claims with financial assets absorbed 30% of GNP, I for one would prefer fewer assets than claims. This point is also suggested in the Kihlstrom-Pauly paper.

The final section of the Brainard-Dolbear paper is both ingenious and incomplete. Proposals which attempt (1) to apply to the micro-economic expected utility apparatus in a general equilibrium context, (2) to permit assignment of laborers to other employers (seemingly a form of slave trade), and (3) to establish futures markets in human capital are heady stuff. Their analysis certainly requires a much more detailed model than appears in this paper.

The paper by Kihlstrom and Pauly, on the other hand, appears to be cautiously executed and very suggestive. For a highly-simplified, one-commodity model, they show that equilibrium in their insurance market is equivalent to a contingent-claims, competitive equilibrium. They also consider whether or not the set of available assets—in this case insurance policies—is sufficient to achieve Pareto optimality. In their model, the number of necessary insurance policies is less than the number of contingent claims, but nevertheless large. Drawing on results of Radner, they report a very interesting discussion about how to evaluate market efficiency when incomplete information exists. The discussion notes the public-good character of information and perhaps for this reason the last half of the paper is somewhat less rigorous than the first half. This remark is not meant as criticism, for the treatment of public goods in general equilibrium models is in its infancy. Kihlstrom and Pauly have written a very stimulating paper.

POPULATION AND ENVIRONMENT IN THE UNITED STATES

Impact of Population on Resources and the Environment

By JOSEPH L. FISHER
Resources for the Future, Inc.

Without any question one of the liveliest subjects before the American people during the last year or two has been population, natural resources, and the environment. The subject is not new to economists. The classical English economists, especially Malthus, were concerned about the effects of population growth on land and agricultural production. But an important new element has appeared quite recently, particularly in this country and in other more developed countries, which is leading economists to reexamine the population and environment question. That is the visible deterioration in the quality of the environment.

Resources may be regarded in two lights: resources as commodities and resources as environment. Malthus and the other greats were concerned mainly with the first, resources as commodities. The second aspect, resources as environment, has only recently emerged as a major topic of consideration by economists. By environment one means, in the most inclusive sense, the entire biosphere of the earth. More particularly one has in mind the basic natural resources of land in its several uses—fresh water and oceans, minerals in place, the atmosphere in its several layers, and the living and working space on the surface of the earth. One also is thinking about cities and metropolitan regions, river basins, the countryside, and other geographic areas.

The purpose of this brief paper is to examine the principal impacts of population growth in the United States, and its distribution, on resource commodity availability and on environmental quality. Most of the points made apply to other more developed countries (much of Europe, Japan, Australia, and a few others), but have to be modified considerably with reference to less developed countries.

Resource Effects

Looking first at the impact of population growth in the United States on resource commodities, the historical record of the last 100 years does not indicate any general supply stringency at least for the next few decades. For all resources (agricultural, forest, and minerals) the average increase in per capita consumption has been about one-third percent per year, with virtually no increase in recent years for agricultural and forest products, but with continued increases for mineral products including fuels. Judged on this basis then, the resource scarcity has decreased persistently for a long period of time.

Another indicator of resource availability is the trend of the employment/output ratio for the resource industries. Here the record shows persistent and considerable decreases, with the exception of forest products and possibly fish, among the more significant raw materials. Less than 10 percent of the nation's labor force, in-

stead of the more than 50 percent in 1870, now produces over 20 percent more in resource products per capita. Employment/output statistics point to diminished rather than increased scarcity; fewer workers working much shorter hours now produce much more than was turned out a hundred years ago.¹

If resource commodities were becoming scarcer in the economic sense, one would expect their costs and prices relative to costs and prices generally to have increased. Raw material price data over the past century, as best they may be pieced together, do not exhibit any marked tendency to rise or fall except for the upward movement of forest products. Instead they have moved erratically up and down as a result of business cycle factors, world supply and trade conditions, and applications of new technology, although in the past few years a case could be made for a general upward movement. Thus, for this most obvious and direct indicator of resource availability, the movement of relative costs and prices, the overall picture does not indicate increasing scarcity, certainly on any general or alarming scale.

It has to be noted, however, that the United States now relies quite heavily on foreign sources for a number of important raw materials such as oil, iron ore, bauxite, copper, lead, zinc, and numerous other metals. For a few items, imports constitute virtually the sole source of supply for this country. In recent years imports of crude oil have been around 20 percent of consumption, imports of iron ore around 30 percent. Without these large imports, costs and prices of these two exceedingly important products undoubtedly would have risen on the basis of further exploita-

tion of leaner and deeper domestic sources. For a number of the major agricultural commodities, the United States remains a net exporter. But a case can be made on the basis of increased imports of a growing number of basic raw materials over the last few decades that, viewed in national self-sufficiency terms, the resource situation in this country has become less favorable.

A final indicator of resource commodity availability has reference to the reserves and potentials for meeting requirements. By the end of the century, with perhaps a 50 percent increase in population to around 300 million and perhaps a three-fold increase in gross national product to around \$3,000 billion in 1970 prices, it is perfectly obvious that raw material inputs into the American economy in the year 2000 will be much greater than they are now. As rough estimates the following may be contemplated: total energy requirements, up one and a half to two times by 2000 over 1970; electric power, up by six times or so; iron ore, more than double; aluminum, up by five times and perhaps much more; copper, double or more; timber, more than double; fresh water withdrawn and not returned, up one and a half times; wheat, up by one-quarter to one-third; cotton, up one and one-half times or more.²

It becomes a complicated task to check requirements of this order against reserves and potentials; one does not know that additional sources will be discovered in the coming years nor can one predict the course of technological development from which can come substitute materials, new products, and more efficient processes. Nor can one see clearly new forms of management or new public policies which could greatly stimulate (or conceiv-

¹ Joseph L. Fisher and Neal Potter, "Natural Resource Adequacy for the United States and the World," in *The Population Dilemma* (second edition), ed. by Philip M. Hauser (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969).

² These rough estimates are based on statistics and work in progress at Resources for the Future.

ably repress) the resource industries. Yet, for what it is worth, a few observations may be offered regarding future supply possibilities.

For a number of decades past, most of the food requirements of the nation have been met without increasing the number of acres of cropland; this trend will probably continue at least for a while into the future and may even hold for the balance of this century depending upon the rate at which improvements continue to be made in American agriculture. Overall, national availability of fresh water has little meaning; one has to examine the situation in terms of river basins or portions of river basins. Adequate supplies of fresh water undoubtedly will not come easily in a number of places, particularly if high quality water is required. The main problem for water resources will be adequate planning, investment, and management. For many of the metals the only protection against steep increases of costs and prices will be enlarged imports, although new technology for recovery, smelting, and refinement would prove helpful, as would increased substitution of plastics, glass, and some other materials. For the all-important energy commodities the outlines for supply during the next three decades appear to be as follows: plenty of coal; enough oil with perhaps some tendency for price increase later on depending on imports, development of new supplies in Alaska and elsewhere, technological improvements in recovery of liquid petroleum products from oil shale and coal, and above all the rate at which nuclear power comes into the economy; and less natural gas, especially from domestic fields, unless pipeline deliveries from Canada and Alaska plus shipments in liquid form from the Middle East, North Africa and perhaps elsewhere become much larger. The big hope for the more distant future lies in nuclear power, first breeder reactors and ultimately fu-

sion, which though promising to ease greatly any supply problems would create major environmental difficulties.

Overall the outlook for the United States economy for resource commodities can be guardedly optimistic. Past trends and future indications point to general adequacy in some sense, though with numerous cost-price and supply difficulties for particular things, in particular places, and at particular times. This favorable outlook, it must be noted, is predicated on several important points: the continuation of a reasonably free international trading and investing system so that the United States can continue to draw upon foreign sources for a number of resource materials, of scientific and technological progress translated into economic production so that most incipient shortages can be circumvented by resort to new sources and substitute materials, and of improvements in policy.³ Failure at any of these points could quickly remove any general optimism of outlook. A series of deep and protracted depressions would, of course, greatly alter the outlook, in many ways in a favorable direction so far as a resource availability is concerned; and in case of a major war all bets are off. The view beyond 2000 is clouded with uncertainties about population, technology, demand patterns, and styles of living. Many of the possible scenarios that shimmer in the crystal ball are frightful. These are warnings that should stimulate preventive action and not paralyze it.

One other condition will have to be satisfied for a favorable outlook regarding resource commodities to hold true, and that has now assumed a towering importance. The quality of the environment has to be

*These conclusions are drawn from Hans H. Landsberg, Leonard L. Fischman, and Joseph L. Fisher, *Resources in America's Future* (Baltimore: The Johns Hopkins Press, 1963) and from subsequent articles stemming from the research.

assured, at least at some minimum level below which neither population nor economic growth can be sustained satisfactorily. Let us now briefly consider the impact of growth on the environment.

Environmental Effects

Residuals waste materials that make up the bulk of what we call pollution are associated with the increase of both population and production, especially of particular items. The following are major sources of several types of pollution; they indicate a tremendous increase in environmental pollution generally in this country during recent years. A large fraction of solid waste that has to be disposed of consists of paper and paperboard products. Since 1950 the amount produced in the United States has more than doubled from 24 to 53 million tons. The major contributor to air pollution in this country is the motor vehicle. The number of motor vehicle registrations has more than doubled since 1950 while the total vehicle miles have increased by an even larger percent. Vehicles scrapped, an indicator of another kind of solid waste, rose from 3.2 million in 1950 to 7.3 million in 1967. Heat pollution whether discharged to the atmosphere or to water courses may be indicated by electric energy production which increased by approximately four times between 1950 and 1969. Electric energy is also a good indicator of sulfur dioxide pollution in the atmosphere. Chemical fertilizers which are washed off the land into streams, estuaries and bays, and lakes in many instances result in eutrophication and consequent damage to aquatic life and the use of water for recreation. In the last two decades the use of fertilizer has doubled. During the same period the use of DDT and similar pesticides which persist in the environment has increased phenomenally.

Pollutants that result from these and

other production and consumption activities need not increase in precise ratio with production of these items. The residual materials may be recycled before they escape into the environment, or they may be recovered from the water courses or the landscape and recharged into industrial processes. They may also be reduced in bulk or potency, neutralized, or otherwise abated. To the extent any of these can be done, the potential damages from pollution will be assuaged.

Direct monitoring and measurement of pollutants discharged into the environment would be more useful for many purposes than relying upon indirect indicators such as time series of production and consumption. Unfortunately statistics of this direct kind are not available as yet on a comprehensive scale, at least for more than a scattered few years. One of the early tasks confronting the Council on Environmental Quality which was established in early 1970 is to bring order into the statistics of pollution. Steps also need to be taken to begin more comprehensively to measure the economic, social, and health effects of pollution.

Other types of environmental quality deterioration include traffic congestion, particularly during commuting hours in the metropolitan areas; poor housing in both city and rural slums; noise, especially near airports, highways, and certain industrial operations; and failure to protect and manage nature parks and wilderness areas properly. Many of these may be classified as problems of the amenities, although some of them may lead to more serious consequences.

Special note must be made of certain types of pollution and environmental degradation which are of a more serious order. One type involves possible damage to the genetic stock of human beings, animals or plants. Radiation pollution is the most obvious example. Another type,

while posing no genetic threat, may injure or even kill anyone subjected to an overdose. A dramatic example currently in the news is mercury poisoning. A third profoundly serious type involves possible destruction of ecosystems upon which human beings (and perhaps other or all forms of life) depend for life and sustenance. In this category would fall those pollutants and environmental disturbances that drastically alter climate or the operation of oceanic and atmospheric systems. An example would be furnished by continued buildup of carbon dioxide in the upper atmosphere, although at the present time not enough is known on this subject to draw any conclusions except that more research is needed.

So much for sketching some of the principal impacts of population and economic growth on resources viewed first as commodities and then as environment. How much pollution and environmental degradation is to be accounted for by population increase and how much by growth of production and consumption of goods?

The United States gross national product for the past 25 years has been increasing three or four times as fast in constant dollars as population and very likely will continue to do so for the remainder of this century, and beyond. If GNP continues to increase at about four percent a year (which is not far from what has happened in recent years), then three (or somewhat more) of the four percentage points is due to productivity increase and one percentage point (or slightly less) to population increase. According to this way of looking at it, one could conclude that productivity increase outweighs population increase by at least three to one as a cause (or explanation) of economic growth generally and concomitantly as a cause of waste residuals.

For specific industries, technology and economic growth are overwhelmingly the

cause of pollution. For example, 90 percent of the growth in power generation in the last 30 years has been caused by higher per capita consumption and only 10 percent by population growth.⁴

As far as most pollutants are concerned, national population and production figures do not reveal the full story by any means. One has to know also a good deal about the distribution of population: how many live in metropolitan regions, and within them in central cities, and how many in nonmetropolitan areas. During the last 20 years population of the standard metropolitan statistical areas has increased nearly 50 percent, from 89 to 129 million, and now makes up nearly two-thirds of the national total. Within the SMSA's, population of the central cities has increased very little while the population outside the central cities has virtually doubled. The tendency toward further concentration in metropolitan regions outside central cities seems certain to continue during the several decades ahead though perhaps not at as rapid a rate as during recent years. All this means that airshed, watershed, and urban solid waste kinds of pollution will be magnified in the future as population concentration continues to build up in metropolitan areas. The same will be true regarding people and vehicle congestion, the overuse of parks and public facilities, noise, and some other urban-type disturbances. Production facilities, including those which cause pollution, tend to be even more concentrated in metropolitan areas than population and consumption.

Looking behind the economic growth component of both resource consumption and environmental degradation, one must note that not all industries and forms of consumption increase equally, and not all

⁴ Hans H. Landsberg, "A Disposable Feast," in *Resources*, published by Resources for the Future, Inc., June 1970.

of those that increase have harmful environmental effects. The principal contributors to water pollution are no more than half a dozen major industries such as pulp and paper, food processing, and several others. The principal contributors to air pollution are the automobile-gasoline complex, electric generating plants, industries that produce burnable solid wastes, and a few others. Clues can be found here as to how to focus research efforts to reduce pollution through such measures as recycling, product and process redesign, pollution abatement devices, and so on.

Of course underlying and giving impetus and direction to production is the consuming power and preferences of people. If consumers would turn their attention away from pollution-associated products and processes, or would be willing to pay the extra cost in time or money of preventing or abating pollution, improving environmental quality would be much easier. Thoughts along this line lead naturally then to a consideration of the value system which underlies economic behavior in this country. This is widely regarded as primarily materialistic in nature and oriented to economic growth and more consumption. Changing underlying values is a difficult and serious matter beyond the reach of economics strictly defined.

Recently a number of economists are trying to reshape economic theory and analytical tools so as to take rationally into account cost and benefit externalities, the importance of which has lately become widely apparent. The central thrust here is to bring within the ambit of ordinary private and public economic decision-making external effects that until recently have been regarded as external to these decision-making units and therefore not taken into account in any responsible fashion on behalf of society. The management of air, water, and other common

property resources, frequently on a regional scale, has been at the center of their attention.⁵

The outlook for resources as environment appears less certain and more charged with problems, even dangers, than does that for resources as commodities. Much will depend on how successfully new policies can be applied, the extent to which behavior about the environment can be changed, and how ready people are to experiment with new institutions for environmental management. With a diversified technology and rich economy, a well-trained labor force and skillful managers of public and private enterprises, and, one hopes, a will to do better on the part of everyone, this nation should be able to bring its environmental problems under control. The problems that are of international scope will require equally vigorous attention.

Toward a Population-Resource-Environmental Policy

What is lacking, and much needed, is a more coherent set of policies in this field. The nation should move toward that population, and population distribution, most consistent with (1) fostering improvements in material and nonmaterial levels of living for all, and (2) protecting human life and the sustaining capacity of the biosphere indefinitely into the future. These general objectives will have to be translated into specific policies and actions, among other things to stabilize fertility rates, change the distribution of people and jobs, develop new and substitute resource materials that are cheaper and less polluting, recycle residual wastes, establish and enforce minimum environmental

⁵ For an imaginative approach along these lines see Allen V. Kneese, Robert U. Ayres, and Ralph C. d'Arge, *Economics and the Environment: A Materials Balance Approach* (Washington: Resource for the Future, 1970).

quality standards and provide incentives for higher standards, gain more information and knowledge about population-resource-environmental problems, and create or remake institutions for management and control.

Many of these will have to be carried out by means of economic measures because it is the producing and consuming behavior of people and organizations that is to be altered. At the broadest level of policy, the injunction in the Employment Act of 1946 that government promote "maximum employment, production, and

purchasing power" needs to be combined effectively with the injunction in the National Environmental Policy Act of 1969 to promote "efforts which will prevent or eliminate damage to the environment and biosphere and stimulate the health and welfare of man." Many of the population-resource-environmental issues of the 70's will turn on finding satisfactory reconciliation between production and consumption of resource commodities on the one hand, and protection of the environment on the other.

Does Human Fertility Adjust to the Environment?*

By RICHARD A. EASTERLIN
University of Pennsylvania

From Malthus to Paul Ehrlich, there have been those who see the growth in man's numbers driving him inexorably, like lemmings to the sea, towards misery and death. A basic premise of this view is that human reproductive behavior does not voluntarily respond to environmental conditions. Instead, man, following his natural instincts, will breed without restraint and population will grow until environmental limits force a halt through higher mortality.

Nowhere, I think, is this view called more into doubt than by American historical experience. Here, if anywhere, environmental constraints on population growth at the start of the nineteenth century appeared to be at a minimum. Certainly this is what Thomas R. Malthus himself thought. Writing of "the English North American colonies, now the powerful people of the United States of America," he said:

To the plenty of good land which they possessed in common with the Spanish and Portuguese settlements, they added a greater degree of liberty and equality. Though not without some restrictions on their foreign commerce, they were allowed a perfect liberty of managing their own internal affairs. The political institutions that prevailed were favourable to the alienation and division of property. . . . There were no tythes in any of the States, and scarcely any taxes. And on account of

the extreme cheapness of good land, a capital could not be more advantageously employed than in agriculture, which at the same time that it supplies the greatest quantity of healthy work, affords much the most valuable produce to the society.

The consequence of these favourable circumstances united, was a rapidity of increase, probably without parallel in history.¹

Indeed, American fertility was extremely high. According to Coale and Zelnik, the birth rate at the start of the nineteenth century "was markedly higher than that ever recorded for any European country and is equaled in reliably recorded data only by such unusually fertile populations as the Hutterites and the inhabitants of the Cocos-Keeling Islands."² The reproductive performance of a number of today's high-fertility nations is modest by comparison with the early American record.

The astounding thing is that from about 1810 on, American fertility started to decline. And this, shortly after a vast expansion of natural resources had been accomplished through the Louisiana Pur-

¹ Thomas Robert Malthus, *First Essay on Population* 1798 (New York: Augustus M. Kelley, 1965), pp. 104-105.

² Ansley J. Coale and Melvin Zelnik, *New Estimates of Fertility and Population in the United States* (Princeton: Princeton University Press, 1963), p. 35. While source references in this paper have been kept to a minimum, a number are given in Lance E. Davis, Richard A. Easterlin, and William N. Parker, eds., *American Economic Growth: An Economist's History of the United States* (New York: Harper & Row, forthcoming 1971), chapter 5, and Richard A. Easterlin, *Population, Labor Force, and Long Swings in Economic Growth* (New York: Columbia University Press, 1968), Part II.

* Comments on an earlier draft of this paper by Glen G. Cain, Joseph L. Fisher, Albert Fishlow, Ronald Freedman, William H. Newell, and Almarin Phillips are gratefully acknowledged. This research was supported in part by National Science Foundation grant NSF-GS-1563.

chase! While the data are not perfect, by 1860, for the white population, the ratio of children under five to women 20-44 years old (the fertility measure most generally available) had fallen by a third from its 1810 level, and by 1910, by over a half. Put differently, in 1790 almost half of the free families contained five or more persons; by 1900, the proportion of families with five or more persons had fallen to less than a third.

How can one reconcile this dramatic reduction in fertility with the seemingly abundant state of natural resources throughout much of this period? In this paper, I shall suggest that, while our state of knowledge is far from adequate, a plausible case can be made that the secular decline in American fertility was a voluntary response to changing environmental conditions. For obvious reasons we need to know much more about the mechanisms involved, but if the line of reasoning here is correct, it refutes the analogy commonly drawn between human population growth and that of fruit flies in a jar.

Theory

The analysis builds on the economic theory of fertility. In this theory, tastes, prices, and income determine the optimal number of children. The optimal number of children, together with infant and child mortality conditions, determines the optimal number of births. Finally, the extent to which actual births exceed optimal births depends on attitudes toward and extent of information about fertility control practices, and the supply conditions of such practices.

The population is subdivided into several component groups, each subject to rather different conditions. For the present discussion, attention is focussed on a classification in terms of location—frontier areas, settled agricultural areas, new urban areas, and old urban areas. The ar-

gument is that the basic fertility determinants—in particular, the cost of children, fertility control practices, and factors other than income change influencing tastes—vary among these locations in such a way that fertility tends to be progressively lower as one moves from the first to the fourth of these situations. Since the course of American economic growth in fact involved a population shift in just this direction, the result was a continuing secular pressure toward fertility reduction.

To take up first the matter of tastes, two determinants directly bound up with economic and social development are the progress of education and the introduction and diffusion of new goods. Both tend to alter preferences for goods versus children in a manner adverse to fertility, because they create or strengthen consumption outlets competitive with children as a source of satisfaction. Education creates awareness of new modes of enjoyment and opens access to them. (For example, while children are a recognized source of satisfaction to all, it is usually only persons of higher educational status who consider foreign travel as a serious consumption alternative.) The progress of education during economic development means that a growing number of households experience a widening of consumption alternatives. Considering different locations at a given time, education was typically more advanced, and hence this influence stronger, in older rural areas than on the frontier, in urban areas than rural, and in older urban areas than in newer urban areas.

Much the same type of argument may be made regarding new goods. New products were more available in older rural areas than on the frontier, because the marketing system was more advanced; hence a wider range of items competed with children in the older areas. Similarly, people in urban areas were more exposed

to new goods than people in rural areas by virtue of the greater market potential offered by the denser populations, residents of older urban areas being more exposed than those in newer.

As for costs of childbearing, both the outlays on and returns from children tended to create cost differentials among areas with an effect on fertility similar to that of tastes. On the frontier, with its demands for breaking and clearing new land, the potential labor contribution of children was greater than in established agricultural areas. Also, with land relatively abundant, the problem of establishing mature children on farms of their own was less serious. Nevertheless, in the established agricultural areas, the labor contribution of children on family farms was higher than in cities where work possibilities were more restricted. At the same time the costs of raising children were higher in the cities, since food and housing were typically more expensive there than in rural areas. Thus, taking account of both costs and returns, children were increasingly expensive as the situation changed from frontier to settled agriculture to an urban location.

Finally, consider the situation with regard to methods of fertility limitation. In general, knowledge and availability of a variety of fertility control practices were greater in urban areas than in rural. Similarly these conditions were better in settled agricultural areas than on the frontier, which was at the periphery of the communications network.

Putting together these influences—tastes, cost, and fertility control practices—leads one to expect the following ordering of areas from high to low fertility at any given time: frontier, settled agriculture, new urban areas, old urban areas. Also since frontier areas gradually become transformed into settled agriculture and new urban areas into old, one would ex-

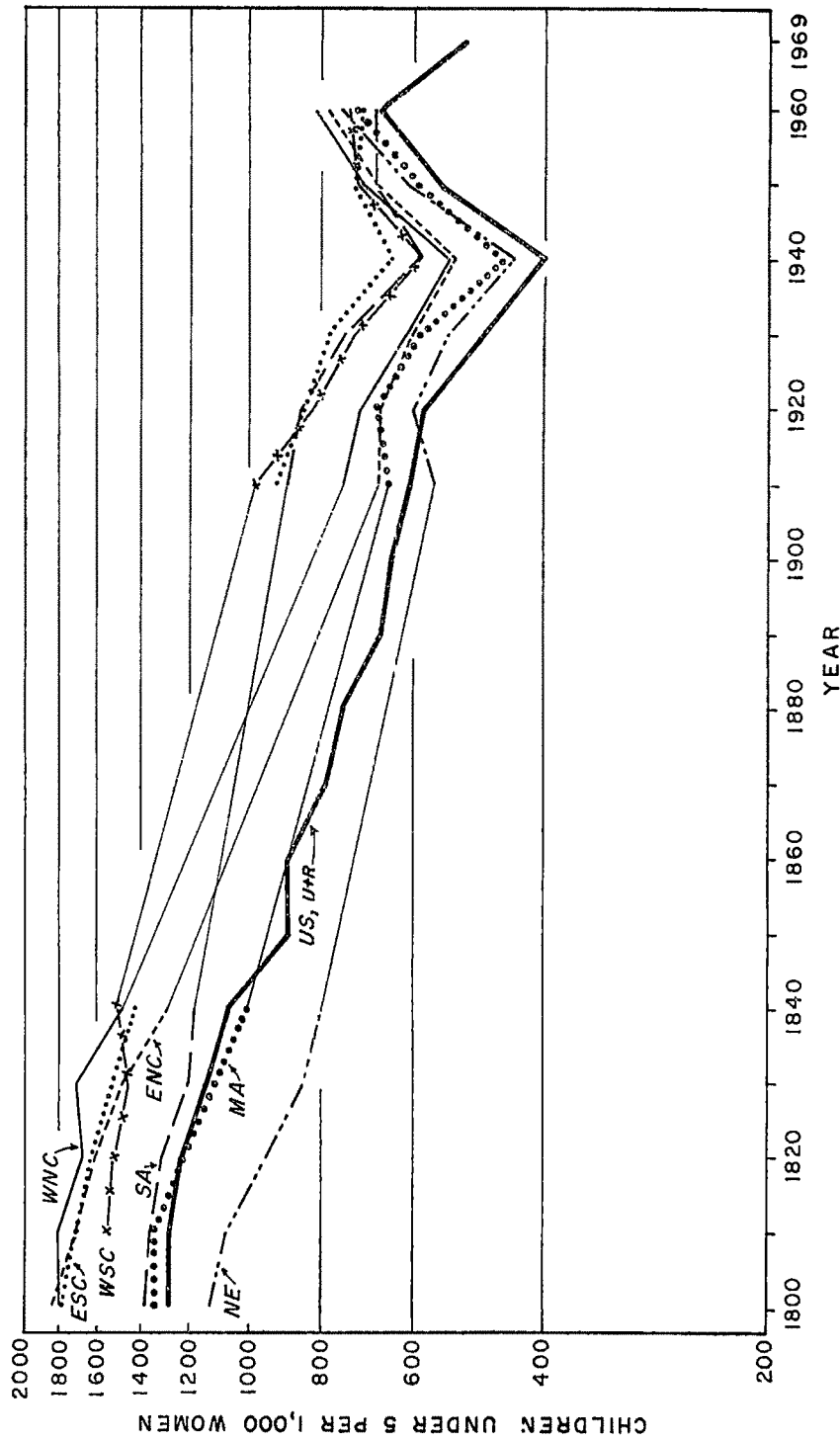
pect that over time fertility would decline as new areas “age”. Moreover, since “new” and “old” are matters of degree, not of kind, one might expect that even in settled agricultural areas and older urban areas, fertility would continue to decline, at least for some time, as “aging” continues.

In urban areas this process of aging is reinforced by the trend in the composition of the population by origin. Initially, the populations in urban areas are dominated by in-migrants from rural areas or abroad who bring with them a high fertility heritage. In the course of time, these first generation urbanites are gradually replaced by second and later generations born and raised in urban areas, and with consequently lower fertility “tastes.” Even today fertility among new rural migrants to urban areas exceeds that of persons who originate in urban areas. Thus to the pattern of cross-section differences by location noted above, one can add an expectation of fertility declines through time in all four locations.

Evidence

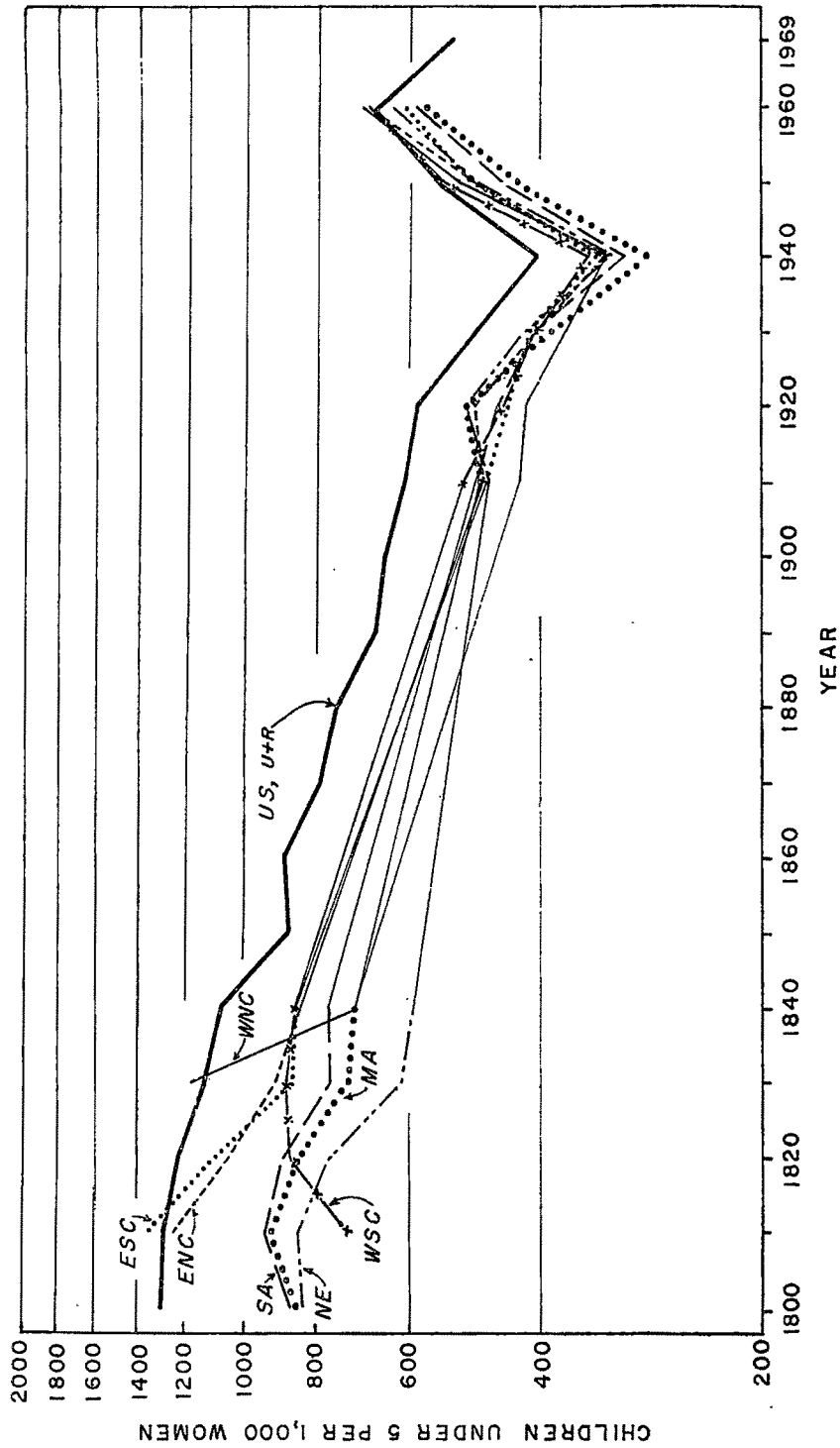
These expectations are supported by the available evidence, though much more work is needed. In Figure 1, the ratio of children under five to females aged 20–44 is shown for the rural white population in each geographic division from 1800 to 1960; Figure 2 presents similar data for the urban white population. In both figures the United States ratio is depicted by a heavy line. This is the average for the population as a whole, rural and urban combined, and therefore provides a common reference point in the two figures. Unfortunately the rural-urban data are not available from 1850 to 1900; hence the 1840 and 1910 observations for each division are connected with a thin line. The Mountain and Pacific divisions have been omitted, because of the absence of data for the period when they were being

FIGURE 1. NUMBER OF CHILDREN UNDER 5 YEARS OLD PER 1,000 WHITE WOMEN 20 TO 44, UNITED STATES, 1800-1969, AND RURAL, BY DIVISION, 1800-1840 AND 1910-1960



Source: U.S. Bureau of the Census, Historical Statistics of the United States, Colonial Times to 1957 (Washington, 1960) and Historical Statistics of the United States, Colonial Times to 1957; Continuation to 1962 and Revisions (Washington, 1965) Series B39-B68. The 1969 U.S. figure was kindly supplied by Wilson H. Grabill of the Bureau of the Census. Urban and rural classified by 1940 Census rules.

FIGURE 2. NUMBER OF CHILDREN UNDER 5 YEARS OLD PER 1,000 WHITE WOMEN 20 TO 44, UNITED STATES, 1800-1969, AND URBAN, 1800-1840 AND 1910-1960



Source: Same as Figure 1.

settled. One would expect these divisions to be somewhat different from the others, however, because of the importance of mining rather than agriculture in their early development. Similarly, the non-white population has been omitted for lack of data. It is the white population, however, which predominantly accounts for the national patterns.

The child-woman ratio used here as a fertility measure typically exceeds the crude birth rate by a factor in the neighborhood of 20 to 25. Analytically, this reflects the fact that the child-woman ratio is computed from (a) a denominator about one-fifth as large as that for the crude birth rate (females aged 20-44 instead of the total population), and (b) a numerator four to five times as large. (Implicitly, birth experience over a five-year period is totalled rather than averaged, and is multiplied by a survival rate on the order of .80 to .95 to exclude those dying before the end of the period.) Although the child-woman ratio is an imperfect measure of fertility, it provides a reasonable basis for a preliminary test of the implications of the foregoing analysis.

Consider first the differentials by location in the early part of the nineteenth century. As shown in Figure 1, among rural areas, fertility was lower in the older settled areas in the East than in those undergoing settlement, the areas west of the Appalachians. This differential between new and old areas existed in both the North and the South. Differences among areas in the age distribution of women of reproductive age, whether due to vital rates or migration, had little to do with this fertility differential. The same differential between newer and older regions holds also for urban areas (Figure 2). Thus it is the three older East Coast divisions which are grouped together at the bottom of Figure 2. Even the seemingly partial exception, the West South Central

division, is not really an exception, because the figures are dominated by Louisiana, an area which was settled early. Finally, within every division, urban fertility is lower than rural fertility. This is quickly seen by comparison with the United States reference line in the two figures. Virtually all of the urban ratios are below the United States average, and except for New England, almost all rural ratios are above. Thus, there is a clear and consistent pattern of fertility in frontier areas exceeding that in older established areas, and of rural fertility generally exceeding urban.

With regard to trends, the United States ratio declines from 1810 on, and this is seen to occur in both rural and urban sections of all geographic divisions, though with some differences in timing. Frontier areas become progressively settled and new urban areas are transformed into old, while within the older rural and urban areas the process of aging continues.

Further evidence consistent with the present interpretation is provided by Yasuba's analysis of the fertility ratios of states for the period 1800-1860. His principal result, which bears particularly on the rural fertility patterns, is that the most important factor associated with fertility differences and trends was population density—the higher the density, the lower the fertility ratio. Yasuba interprets population density as a measure of the availability of land, and argues that:

... [In] a community where the supply of land is limited, the value of children as earning assets is low and hence the demand for children may not be so great as where there is plenty of open land nearby. The increased cost of setting up children as independent farmers and fear of the fragmentation of family farms may further encourage the restriction of family-size in densely populated areas.³

³ Yasukichi Yasuba, *Birth Rates of the White Population in the United States, 1800-1860* (Baltimore, Johns Hopkins Press, 1962), p. 159.

More recently, Colin Forster and G. S. L. Tucker have extended Yasuba's analysis, strengthening and refining the basic finding of inverse association between density and fertility.⁴

Mention should be made of another possible influence in the secular fertility decline, the lowering of infant and child mortality. The child-woman ratio fails for the most part to reflect this influence, since it relates not to births but to surviving children as of the census date. The trend in American mortality for much of the nineteenth century is uncertain. But it is clear that from the late nineteenth century onward, there was a substantial reduction in infant and child mortality, and this probably strengthened the tendency toward fertility decline, since fewer births would be needed to obtain a given number of surviving children.

The question arises whether the growth of per capita income associated with economic development may have exerted a strong counter force tending to raise fertility. The answer to this, in my view, is that income growth is a two-edged sword. On the one hand, it tends to make for higher fertility by augmenting the resources available to a household. On the other, it tends to lower fertility through what might be described as an "intergenerational taste effect." The argument in its simplest form is that in a steadily growing economy successive generations are raised in increasingly affluent households and hence develop successively higher living aspirations in the course of their normal upbringing. Thus, while, on the one hand, each generation on reaching adulthood normally has more resources at its command, on the other, it has greater goods aspirations. If long-term growth is not steady, but fluctuating, temporary dispar-

ities can result between the growth of resources and that of aspirations with consequent swings in fertility, such as the United States has been experiencing in recent decades. Secularly, however, the two influences tend to cancel out in their effect on fertility. Whether they do so completely remains an open and important empirical question. It is clear, however, that if this view is adopted, then the presumption no longer holds that secular per capita income growth tends to raise fertility—the net effect could be positive, negative, or zero.

Another question relates to the possible impact of immigration on native white fertility. The first president of the American Economic Association, Francis Amasa Walker, argued vigorously in an article entitled "Immigration and Degradation" that "as the foreigners began to come in larger numbers, the native population more and more withheld their own increase."⁵ In Walker's view:

The American shrank from the industrial competition thus thrust upon him. He was unwilling himself to engage in the lowest kind of day labor with these new elements of the population; he was even more unwilling to bring sons and daughters into the world to enter into that competition.⁶

The argument that immigration led to lower native fertility is suspect, however, because declines in American fertility occurred in times and places not only where immigration was high, but also where immigration offered little or no competition. Thus, as has been noted, fertility turned down early in the nineteenth century, before any substantial influx of immigrants occurred. Recent studies in colonial demography suggest that in parts of New England such declines occurred even in the eighteenth century. Moreover, fertility

⁴ Colin Forster and G. S. L. Tucker, *Economic Opportunity and American Fertility Ratios, 1800-1860* (forthcoming).

⁵ Francis A. Walker, "Immigration and Degradation," *The Forum*, Vol. 11 (August 1891), p. 638.

⁶ *Ibid.*, p. 641.

declines occurred not only in the areas of immigration but in others as well, notably the American South. What seems to be a common feature of many areas where fertility declines set in is not that they were experiencing competition from foreign immigrants, but that the land supply had been largely exhausted and the process of settlement completed.

At issue here is the likely course of the American birth rate had there been no immigration. If one follows the line of reasoning suggested previously, an important consideration is the type of economic activity that the native population would have followed. In the absence of a foreign labor supply in the nineteenth century, it seems plausible to suppose that the workers for American industrial expansion would have been drawn more from rural areas than was actually the case. Correspondingly, the rate of settlement would have been slowed. As a result, over time the native-born population would have been less involved in new settlement and more engaged in urban activities, that is, more exposed to a low than a high fertility environment. This implies that if there had been no immigration, not only is it doubtful that native fertility would have been higher than that actually observed, but it might possibly have been lower.

It is worth noting that the fertility declines of the past were accomplished entirely by voluntary action on the part of the population. To some extent marriage was deferred. But also there were declines in fertility within marriage. These developments took place in a situation where not only was there no public policy to help those interested in fertility limitation, but attitudes, and even laws in many states, were hostile to the practice or even discussion of contraception or other fertility control practices. To emphasize the voluntary nature of this development, however, is not to suggest that there was no need

for family planning policies then or, for that matter, that there is no need today. What historical experience shows is a voluntary decline in fertility, but this is not necessarily the optimal rate of decline from the point of view of social welfare. One can only conjecture how many households suffered from the miseries of unwanted children because of the hostile public environment. Today our evidence on this matter and on the need for intelligent public policy is straightforward.

In the literature on the demographic transition, the secular fertility decline is typically linked to the processes of urbanization and industrialization. But the American experience suggests an additional dimension associated with the transformation of a rural area from frontier to settled agriculture. Nor is this peculiar to the United States. Canada seems to show a similar pattern, while in Europe the association of high fertility and settlement has been noted with regard to parts of Scandinavia, Finland, and Russia in the nineteenth century. Thus, seen very broadly, the American fertility decline reflects not only the processes of urbanization and industrialization but that of settlement as well. This aspect is crucial, because of its implications for today's less developed countries. The traditional view of secular fertility decline has led to emphasis on industrialization in these areas as a prerequisite to a decline. The American experience raises the possibility that, as population increasingly presses against land resources, fertility declines set in within the rural sector itself.

One may, of course, discount the possibility that American experience is relevant to today's less developed countries. This would carry the highly questionable implication that Americans are in some way a distinctive and exceptionally rational species of mankind. Moreover, the historical experience of Europe and Japan shows

that elsewhere in the world, as similar pressures for reducing population growth have mounted, similar responses have taken place. It seems reasonable to suppose that history will, in time, record that the same was true of today's less developed nations as well. Indeed, fertility declines are already occurring in some of these areas.

Conclusion

The conclusion to which this discussion points is that both theory and the empirical research done so far on historical American fertility suggest that human fertility responds voluntarily to environmental conditions. If this is so—and it seems hard to ignore the evidence—then the nature of what is called “the population problem” takes on a radically different guise. The question is not one of human beings breeding themselves into growing misery. Rather, the problem is whether the voluntary response of fertility to environmental pressures results in a socially optimal adjustment. In thinking about this, it seems useful to distinguish between

the potential for population adjustment and the actual degree of adjustment. The staggering change in American reproductive behavior over the past century and a half clearly demonstrates the immense potential for adjustment. Whether, currently, the degree of adjustment is socially optimal remains a matter for research. Such research requires clarification of the mechanisms through which environmental pressures influence reproductive behavior—a subject to which economic analysis can contribute much more than it has so far. One may add—though this lies outside the scope of the present paper—that further research also requires a deeper understanding of the wants and aspirations of individual families in regard to human welfare. For judgments on the social optimality of adjustment obviously call for a welfare function that takes account of the pervasive human desire to have and raise children. It is to be hoped that such research will go forward promptly, in an atmosphere not of hysterical urgency, but of sensitive concern.

Issues in the Economics of a Population Policy for the United States*

By GLEN G. CAIN
University of Wisconsin

Economists approach the question of determining the optimum quantity of the production of a good by examining how the good's social benefits—the sum of all individual benefits—compare with its social costs. Policies directed toward change in the amounts of production are, of course, examined in terms of the marginal changes in benefits and costs. So it should be with people, or, more specifically, with children. It is easy to state, but difficult to implement, the requirements for an optimal population policy. First, all restrictions on the use of existing family planning methods should be removed to permit the exercise of rationality and free choices by parents. Second, parents should be responsible for the total costs (external as well as private costs) of children they bear, although the costs could sometimes be reduced by means of child-related subsidies directed toward correcting for distributional inequities. The second point is beset by contradictions that arise when distributional transfers, which contain implicit price distortions, are overlaid on an undistorted price structure which marginally equates total costs and benefits (aside from distributional benefits).

The problem of achieving distributional equity is, moreover, only one of the conceptual and empirical problems which a benefit-cost approach to population policy encounters. As in every important appli-

cation of this approach, the analyst must cope with faulty prices, measuring externalities, evaluating goods over time, assessing the probabilities of misestimation of the effects of the applied policies, and with weighing intruding value judgments. These problems may prove insuperable in applying economic analysis to the question of determining the optimum production of people. We won't know until we have tried, and it would be an unhappy commentary on economics if the most important, the most time consuming, and the most costly good produced by households was beyond our ability to analyze in terms of social welfare.

Several assumptions which underlie my economic analysis of population policy may be briefly stated. It is assumed that human choices are involved in the determination of population size and that economic variables—prices and incomes—affect the choices. I rely on the conventional assumption of rationality in household decision-making. This does not assume that markets—economic or political—are organized in such a way as to transform household decisions to socially optimal results. The rationality assumption here simply means that responses are made to price changes for purposes of improving one's utility (or one's family's utility). Money per capita income—even corrected for pollution—should not, however, be the sole measure of utility. In the model of household decision-making presented below, children are viewed as a good yielding utility or psychic income to the parents—a view expressed to econo-

*I wish to express my gratitude to numerous colleagues at the University of Wisconsin, Princeton University, the National Bureau of Economic Research, and elsewhere, for their helpful comments on this paper.

mists most cogently by Becker [2] and to demographers and sociologists in lucid terms by Ryder [12]. (There may be a need to reach an audience of biologists.) The social optimum in population size, should allow for this psychic income of children. It is important to point out, however, that this view of children is consistent with the assumption that the utility function of a parent includes as arguments the welfare of his children, both during the parent's life and whenever the children are acting as adult decision-makers. In this regard children differ from other goods. I do not assume altruism regarding nonfamily members in the utility function of parents.

Let me now specify some of the demographic assumptions which provide the context for the analysis. These assumptions are not intended as predictions, but they do delimit empirical boundaries within which I intend the analysis to be relevant.

Demographic and Technological Background

1. Population growth in advanced economies like the United States depends primarily on fertility. Mortality is of negligible consequence. Migration, which in recent years has contributed from about 10 to 20 percent of the annual growth in population in the United States, is relatively easy to control by governmental policy. The search for a population policy is, then, mainly a search for a fertility policy.

2. The concept of optimality of population size must eventually provide for a rate of population growth that averages out to be zero. Humanitarian considerations dictate that this must be achieved by a lower birth rate. But the statement that zero population growth (ZPG) is an eventual logical necessity (in a closed economy), is less important to this paper than is the empirical judgment (backed

up, I admit, by no evidence) that negative externalities of congestion are likely to be felt as the population grows and, in particular, if the population of the United States were to double or quadruple in size over the next 100 years. (See point 5 below.)

3. Population growth is, fortunately, a gradual phenomenon and is, in fact, subject to reversal. For this reason an assessment and evaluation can be made by each generation of the prospect for congestion and other costs associated with population growth. The metaphor of a bomb does not seem apt, and we should feel less constrained (although not unconstrained) in our customary uses of marginal analysis than we might if explosive discontinuities were present.

4. Although a zero rate of population growth is the only equilibrium rate that can be sustained, no cohort of American women ever had as few as 2.11 children per completed family, which is currently required for replacement or ZPG.¹ The lowest number in our history, 2.27, was experienced by the women who were born in the period, 1906 to 1910, and who reached the ages of peak childbearing in the depression. The highest in this century is a projected 3.36 for women born in 1926-1930 [12]. While it is true that the annual birth rates of the past few years are almost as low as they were during the 1930's, these rates obviously do not measure completed cohort fertility, and we will not know the latter for at least 15 years.

The alarm sound by Kingsley Davis [7] and by Judith Blake [3] in papers containing empirical evidence is that Americans desire more than three children per fam-

¹The number of children per completed family which is currently required for a U. S. population that is just replacing itself, which is the sense in which I use the expression "ZPG," is 2.11 per woman, which amounts to about 2.22 per woman ever married and 2.39 per ever married mother [10].

ily. In an important paper which moderates this alarm, Bumpass and Westoff argue, on the basis of improved techniques for determining wanted and unwanted children, that the desired number of children per family is about 2.5 [5]. A remarkable, even shocking, finding they report is that about 20 percent of all births and 40 percent of births among poor families were "unwanted". If all births were limited to just the "wanted" births and, in this sense, "perfect contraception" were realized, then the current desires of American women would lead to a birth rate close to that required for ZPG. The Bumpass-Westoff study, combined with the direct evidence of lower birth rates in recent years, strengthens the view that the United States population is not exploding. This does not mean that a lower rate of growth would not be optimal.

5. Completed cohort fertility of 2.45 to 3.11 children implies that the United States population will reach 308 million to 408 million in 50 years; and from 412 million to 839 million in 100 years, not allowing for increases due to immigration or reduced mortality.² Such implied increases in density may entail a negative externality, but the severity of this is probably exaggerated because of the confusion between density generally and distributional congestion. The particular problem of congestion in the major urban areas is surely attributable to a suboptimal distribution of people and inefficiencies in the various patterns of land utilization, transportation modes, and other institutional arrangements—all of which are suscep-

tible to specific remedies which should be made even if population were stationary. Curtailing population growth in the near future is neither necessary nor sufficient to cope with urban congestion.

The same conclusion may be stated even more emphatically regarding the pollution problem. The technological sources of pollution—the fuels, phosphates, mercury, insecticides, and so on—are perfectly apparent and each has a number of specific remedies. Policies to reduce the number of people is a gross and ineffective method of combating pollution.

The implicit judgment being made is that, in a situation in which all births are wanted, a given dollar amount of population control—i.e., that amount which is equivalent to what is sacrificed by having one less child—yields far smaller benefits in reduced pollution (or urban congestion) than will that same dollar amount if it is spent on specific pollution control. It is in this policy-oriented sense that I regard population size as a relatively unimportant source of the problems of urban density and pollution. The benefit-cost calculations admittedly cannot be defended adequately in the absence of an economic model of fertility and empirical evidence. These issues are discussed below.

6. Let me mention briefly, and quickly dismiss, one additional red herring in discussions of population growth. In the book, *Population Bomb*, Ehrlich [8] argues that the most serious threat to the survival of civilization stems from the population growth of industrialized nations, notably the United States. His argument, which is made by many others, is that the voracious appetite for consumption of the world's well-off peoples, combined with their increasing numbers, will lead to the depletion of resources to an extent that will impoverish mankind. The production side of the ledger is simply ig-

²A synthetic measure of completed cohort fertility is obtained by summing the 1968 age-specific birth rates for females of all childbearing ages. This yields a completed fertility of 2.45 children per cohort. The sum of the average age-specific birth rates for the last nine years, 1960 through 1968, for which data are available yields the 3.11 figure [10]. I am indebted to the computer program provided by Ansley J. Coale and Etienne Van de Walle for the population projections.

nored or misunderstood. For if production by a person (or a family) covers his consumption, where is the source of impoverishment? The answer may be that there is none, or it may be that the economic system is giving us false information about what constitutes consumption and production. In response to this issue, Nordhaus and Tobin have recently calculated a welfare measure of national production, and they find that the conventional measure of GNP is more likely *understating* the trends in per capita welfare [9].

There is widespread criticism of the "ethic of consumption," but there appears also to be a pervasive lack of understanding of what production is: namely, a means of transforming resources from one form to another and enhancing the value of the resources in the process. How we measure value is what economics is all about. Admittedly, however, what we do with our increased production depends on our preferences or values, which, fundamentally, is what economics takes as given. Nevertheless, it seems premature to conclude that we cannot be trusted to use the proceeds of economic growth in production to buy such goods as cleaner air and an end to poverty; yet it is precisely this lack of faith in reason and judgment that underlies the call for a slowdown in growth.

Economic Analysis at the Micro Level

In the older literature on optimum population size, and in much of the contemporary discussion about population in underdeveloped economies, the target variable is per capita income and the framework of analysis is an aggregative macroeconomic model. Population is a denominator, national income a numerator, and there is attention to the effect of population growth on income, to the ratio of the labor force to population, to aggregate savings, and other macroeconomic variables.

Two objections may be raised to employing this approach to the question of optimum population size in a developed economy like that of the United States. A substantive objection is that insufficient recognition is given to the value of children as consumer goods—a label which is intended to imply simply that parents usually choose to have children and that they get pleasure from children. A second and methodological point is that a microeconomic model of the household is preferable to the macro approach because decisions about fertility are made at the micro level. The two points are related: better predictions about fertility behavior would result from a focus on married couples as a household unit, and this focus should allow for the fact that parents are willing to give up other goods to obtain children. From the standpoint of normative economics, the micro approach is useful because we often want to measure social welfare by summing the welfare of individuals.

An economic model of fertility decisions by households is basically a model of demand for consumer durables. Married couples are assumed to make decisions which aim at maximizing their collective welfare. Children yield utility to parents in the form of psychic income over a span of many years. The household's demand for children is a function of the income of the household, its tastes, the price of children, prices of complementary and substitutable goods, and the prevailing institutional or cultural environment.³

³ Separate considerations of "institutions" and "culture"—difficult terms to define, hence the quotation marks—is arbitrary and will depend on the convenience of the analyst. Generally, I prefer to specify how changes in institutions and culture affect tastes and prices, and in this way take them into account. For example, laws liberalizing abortion can be viewed as lowering its price, both directly and by eliminating the costs of crime and punishment, and as shifting tastes toward making abortion more acceptable.

Prices

The costs of children to the parents consist of the goods and time foregone when bearing and raising children. The "direct" price is mainly contained in the prices of goods and services that are complementary with children: obstetrical services, subsequent medical services, food, housing, education, and so on. The Institute of Life Insurance [13] has recently estimated the direct costs of a child up to age 18 in a four-person family earning \$9,000 annually. By my calculations, the present value of these costs at time of birth is about \$13,300, using an 8 percent rate of discount.⁴

The indirect costs consist of the time parents spend on children, although a careful distinction should be made between leisure time devoted to children and work time given to them. To simplify the presentation of the economic model of fertility, let us assume that the work time involved in bringing up children is entirely borne by the mother. She sacrifices time which could otherwise be spent in market work or leisure. The marginal value of this time may be represented by the wage available to her in the market.

I have estimated that the marginal decision to have a first child entails a loss of market earnings which amount to a present value at birth of about \$11,500.⁵ In

addition there is time spent in homework which I have assumed to require 14 hours per week for a child up to age 3, 10 hours between the ages of 3 and 6, and 5 hours of homework thereafter up to age 14. By evaluating this time at three-fourths the going market wage (to allow for the possibility that discontinuities, corner solutions, imperfections, and the like might keep the imputed home wage below the prevailing market wage), about \$6,100 is added to the present value of the indirect costs.

The total costs to the parents of a child, including the dominant component of time foregone, is around \$31,000. Two points may be made. One is that the willingness to assume these costs by parents offers convincing evidence for the proposition that children provide them with a great deal of utility. The second point is that there are a number of prices included in the overall price of a child, and each component price is potentially changeable by means of a wide variety of policy actions.

Some knowledge about the magnitude of the effects on fertility of changing these prices is, of course, crucial to wise policy decisions, and I will return to this issue later. Suffice to say at this point that the single most important price associated with children appears to be the price of the mother's time. An increase in wage rates, fringe benefits, or some other form of the remuneration of work for women can raise the price of children substantially and would be potentially an important antinatalist policy. If, however, free day care were provided, the price effect of children on the mother's time falls to near zero. Not only would the effect of market work options be nullified by free day care, but also the costs of time in the homework component of child care would be drasti-

⁴These calculations are shown in a larger version of this paper, which is available from the Institute for Research on Poverty, University of Wisconsin. The costs do not allow for the costs of a college education. If the probability of going to college for four years is one-half, and we assume that the cost to the parents is \$1,000 per year, the expected present value of the costs of raising children increases by \$400 to around \$13,700. No allowance is made for earnings by the child, but this may be justified by assuming the child keeps all his earnings. My calculations are crude, but I believe they would be relatively insensitive to the many arbitrary decisions involved.

⁵For measures of the quantity of work time lost I have relied upon the rich detail available in the research of Bowen and Finegan [4]. The wage foregone was assumed to be \$2.00 per hour. These com-

putations are shown in the appendix to the longer version of the paper mentioned in the previous footnote.

cally reduced. If the price of children has any significant effect on demand, free day care ought to be strongly pronatalist.

Income

The demand for children, viewed as a consumer durable, would be presumed to be positively related to income, except for the unlikely case in which children were an inferior good. The qualifications that income should affect expenditures on children more than the quantity of children, and that the relative, in addition to the absolute, change in income is operative, are not likely to lead to a reversal of the proposition that the net effect of income is positive. Of course the effect of income on the demand for children should be measured net of any embodied price effects. Income transfer payments that depend on family size, like family allowance payments, carry price effects in an obvious way. Also, as already noted, price effects are included in those income changes that result from changes in wage rates—especially of wives, and such price effects would be expected to be negatively related to time-intensive commodities like children. In the United States we may presume that wage rates of husbands and of children are negligible influences on the price of children; husbands contribute little to homework and childrens' earnings are an insubstantial part of the lifetime income of parents.

I conclude that the effect of income on the quantity of children demanded is positive but probably small.⁶ The implication

is that any increase in population and depletion of resources which lowers per capita income in time period t will reduce fertility rates in time period $t + 1$, assuming free choice by parents. On the other hand, if per capita income continues to increase, as most of us expect, then fertility will rise, *ceteris paribus*. (But note that the female wage rate will also be increasing.)

Institutional, Cultural, and Technological Constraints

Among the "givens" in an economic analysis of demand functions are: (a) the structure of tastes and preferences, which is related to (b) laws and governmental policies, and, finally, (c) the state of technology. All factors are, of course, to some degree, mutable. Detecting and interpreting trends in people's preferences and fashions is difficult, especially identifying changes in tastes that are causally prior, and not merely responses, to price and income changes. It is interesting to ponder, for example, how much of women's liberation and demands for improved birth control measures stem from the heightened value of women's time represented by higher wage rates. Whatever the source, there appears to be an irreversible trend toward more liberal laws and church policies toward birth control, including abortion. The ideal of "every child a wanted child" is closer to attainment, and this implies a dampening effect on fertility rates, as the Bumpass-Westoff paper makes clear. It also implies that economic theories, which rest on rational choice, should achieve greater explanatory power.

One type of governmental action which casts a shadow on the generally optimistic view sketched up to now is the continued emphasis on war as an instrument of national policy. Modern technology of warfare has made population size less important, but not irrelevant, to the capacity of a nation to wage war. A related motiva-

⁶In research by Cain and Weininger, using cross-sectional observations of fertility rates in standard statistical metropolitan areas in 1960 and 1940, the effect of income on fertility was found to be positive but small—an elasticity of around .2 or .3. The wage rate of the wife was negative in its effect—the "pure substitution" elasticity was $-.4$ to $-.5$. (6) Some inadequacies in the data for this research, shared with all research on the economics of fertility, are mentioned in the last section of this paper.

tion is the drive for ethnic or racial domination. These motives, which presumably reflect a summation of individual values, are likely to produce pronatalist policies that would be uneconomic on grounds other than the alleged national survival. Indeed, the existence of such motives challenges (but does not, I hope, overturn) the assumptions of rationality on which the economic model is based.

The role of technology often tends to dominate the debate on the population bomb, but I will not add to my brief earlier comments. Others on the panel are better equipped to discuss the role of technology in augmenting resources and shaping the environment. Regarding the technology of birth control, parturition, and child care, one would guess that the net effects would be antinatalist: improvements in contraception and the potential ability to choose the sex of children ought to outweigh the pronatalist effects of reductions in subfecundity, infant mortality, birth defects, maternal mortality and morbidity, the discomforts in pregnancy and childbirth, and the onerousness of homework tasks in child care.

Externalities and Distributional Issues

Children are now subsidized in many ways, but, except for families on welfare, the subsidies are indirect; and except for public support of education, the subsidies do not sum to a large fraction of the total costs assumed by parents. Although arguments are made for the external benefits of elementary and secondary education and the research component of higher education, I suggest that parents could assume a much greater share of these costs without violence to equity or efficiency criteria. Home ownership is subsidized through the tax system and loan subsidies, but there are also implicit taxes imposed on this industry by building codes and labor unions. Other items, like medical, rec-

reational, and day-care services may be mentioned, but these are probably not areas in which private costs are pervasively and significantly supplemented by public subsidies, except for programs dealing with a fraction of the poverty population.

Existing and proposed systems of income maintenance are another matter. All plans are family-size as well as income conditioned, and this raises important questions for an optimal population policy: what effects do these children's allowances have on the birth rate, what costs are generated, and what amount of benefits from alleviating poverty offset the costs? We have little empirical knowledge about any of these issues.

These plans lower the direct price of children and, because they impose marginal taxes on earnings at rates of 50 to 100 percent, they lower the indirect price by reducing the effective wage rate facing the wife. The pronatal effects may be moderated in two general ways. One is in the design of the income maintenance program. Thus, to reduce a "birth bonus" the income support plan should include childless poor families and could provide graduated payments for a child according to his age. Both measures enhance equity, and the latter is consistent with matching benefits to costs, since older children are more expensive [13]. For the same reasons the payment schedule per child should be reduced as the number of children increase, to allow for the economies of scale.

Another design feature is to substitute cash grants for all income-in-kind subsidies that are pronatalist, like public housing. A symmetrical step is to design the program to minimize the disincentives to work and, in particular, to keep the implicit tax on earnings of the wife low.⁷

⁷ Those familiar with income maintenance systems will recognize a dilemma here. Given a sizeable guar-

A second approach is to subsidize anti-natalist measures to counter the direct subsidies to children. Family planning and labor market services could be provided free, or by means of direct payments if the pronatal effects of children's allowances were seriously large. But let us note that we are not now even providing unrestricted availability of all forms of birth control. Abortion, for example, is still highly restricted and, where available, appears to be priced too high—that is, priced above the "resource-use costs" of the method.

If income maintenance programs lead the poor to have more children than the nonpoor, someone will probably want to examine eugenic issues.⁸ However, it is not clear that children's allowances will have this effect. If the wages available to wives and preferences are, on average, the same, then the presumed positive effect of income will encourage more children by the nonpoor and may offset the positive effects on fertility of the allowance for poor families.

Externalities associated with a growing population appear to me to be all negative but, for the near future, all minor. A positive externality has been claimed because there will be more geniuses and ideas from a larger than from a smaller population, *ceteris paribus*. There will also be more villains, but the crucial weakness in this argument is identical to that regarding den-

sity and pollution; namely, that there are *specific sources* of research, ideas, and knowledge, and it is most efficient to subsidize these directly. In the long run a serious external diseconomy will be the problem of congestion, broadly speaking. The assumption of diminishing returns to labor implies another negative externality, but in the face of secular rising per capita incomes, this appears to me less important. At some point, these problems probably justify subsidies to birth control technology and, possibly, smaller tax deductions allowed for children.

Conclusions and Agenda for Research

The widely discussed issue of the threat of a United States population bomb is, in my opinion, a nonissue. The question of optimality of population size is, on the other hand, always an issue. Indeed, we cannot escape from having a population policy and an implicit benefit-cost calculation associated with that policy. Is it not better to make our benefit-cost analysis explicit? But the economist's approach to optimality is, like the statistician's approach to randomness, one of choosing correct procedures and not one of choosing a correct outcome. We have many options to alter prices (which are, after all, simply guides and incentives) to take account of private and external costs and benefits of children. The externalities of congestion and the price distortions from income maintenance programs, subsidized day care, and other subsidies to children deserve our attention, although they do not require drastic action. A much more important issue, I believe, are the serious obstacles to full parental control over having the number of children actually wanted.

Action to improve the functioning of the markets is, however, severely hampered by the paucity of our empirical knowledge. We have almost no satisfac-

antee for a family (with no other income), a low tax rate on earned income requires a high breakeven point (where the net income transfer is zero). A high breakeven point means that more families are covered by the program and are, therefore, faced with higher tax rates than they otherwise would be. Lower breakeven points and the reduced numbers of families covered by the plan imply, however, even higher tax rates for those who are covered.

⁸ Insofar as eugenics deals with the quality of the population, one might argue that it is a necessary consideration in determining the optimality of a population size. I view this as the topic of a Super Optima and would defer its analysis to others.

tory empirical estimates of elasticities of the quantity of children demanded with respect to income and prices. Our data for research suffers from a number of inadequacies. The data are nonexperimentally generated. Unmeasured tastes variables (i.e., relative preferences for children compared to other goods) may change slightly over time, but they are probably correlated with such relevant variables as family income in cross sections, and this makes it difficult to obtain unbiased measures of the latter variables. The critical dependent variable is completed family size, but the income and price variables are not measured over the entire span of the childbearing period of women. The task of estimating price and income effects on completed fertility is analogous to that of estimating the total number of cars purchased over the lifetime of the cohort, and our current time series and cross-section data are not well-suited to this estimation task.

We have no hard evidence on the effects of the price of housing, education, day care, or other services relevant to children. I can mention only the journalistic evidence of the antinatalist effects of housing restrictions, incentives to female labor, and provision of contraceptive means in the Eastern European and Scandinavian countries and in Japan. It is plausible that Japan's birth rate is also held down by the emphasis on and high private costs of education.

It should be frankly acknowledged, therefore, that the feasibility of attaining a benign rate of population growth, even in the terms of arguments which are conventional in economics, rests on limited scientific evidence. Let me conclude by mentioning some of the qualifications to the argument which are, conventionally, outside of economics.

The assumption of the dominance of nonmalevolent and rational behavior of

man is something of an article of faith. Population growth, unlike warfare, is, however, a gradual process and the result of millions of decentralized decisions. This provides some flexibility in outcomes, but raises other issues. One is whether the welfare of future generations of people and even nonhuman species will be taken into account. This depends on the persuasiveness of the arguments to do so. There is evidence that people behave in the interests of their children and grandchildren, at least, and, indeed the current generation sacrifices consumption to pass on savings which contribute to a rising per capita income of future generations. I cannot get excited about the concern for future generations under conditions of projected growth in per capita income.

Finally, a challenge to rational decision making is posed by a version of the index number problem. The argument is that we might adopt only one style of life as population grows—in particular, high density urban living—and so fully adapt to this way of living that no comparisons are possible. Ardrey's rats, after living in crowded conditions, can end up with a revealed preference for a way of life that we, in our present state of minds, abhor [1]. I believe, however, that a diversity of life styles—of family sizes and living arrangements—remains open to us for the foreseeable future and that we can choose more wisely than rats.

REFERENCES

1. Robert Ardrey, *Social Contract*, Athenaeum, New York, 1970.
2. Gary S. Becker, "An Economic Analysis of Fertility," in *Demographic and Economic Change in Developed Countries*, a conference of the Universities-National Bureau Committee for Economic Research (Princeton: Princeton University Press, 1960), pp. 209-231.
3. Judith Blake, "Population Policy for

- Americans: Is the Government being Misled?", *Science*, Vol. 164, May 2, 1969.
4. William G. Bowen and T. A. Finegan, *The Economics of Labor Force Participation*, Princeton: (Princeton University Press, 1969).
 5. Larry Bumpass and Charles Westoff, "The Perfect Contraceptive Population: Extent and Implications of Unwanted Fertility in the U.S.," *Science*, Vol. 169, September 18, 1970, p. 1177-1182.
 6. Glen G. Cain and Adriana Weininger, "Economic Determinants of Fertility: Results from Cross-Sectional Aggregate Data," available as a Discussion Paper from the Institute for Research on Poverty, University of Wisconsin, Madison.
 7. Kingsley Davis, "Population Policy: Will Current Programs Succeed?", *Science*, Vol. 158, November 10, 1967.
 8. Paul R. Ehrlich, *The Population Bomb*, Ballantine Books, New York, 1968.
 9. William Nordhaus and James Tobin, "Is Growth Obsolete?", an unpublished paper prepared for a National Bureau of Economic Research colloquium, December 10, 1970.
 10. "Population Estimates and Projections," Current Population Reports, No. 448, August 6, 1970, U.S. Department of Commerce, Bureau of the Census.
 11. Norman B. Ryder, "The Emergence of a Modern Fertility Pattern: United States, 1917-66," in S. S. Behrman, L. Corsa and R. Freedman, eds. *Fertility and Family Planning: A World View*, University of Michigan Press, Ann Arbor, 1969, p. 394.
 12. Norman B. Ryder, "Fertility" in *The Study of Population*, Hansen and Duncan, eds., University of Chicago Press, Chicago, 1959, pp. 400-436.
 13. Sara A. Sohn, "The Cost of Raising a Child," Division of Research and Statistics, Institute of Life Insurance, 1970.

DISCUSSION

T. PAUL SCHULTZ: Professor Easterlin proposes that man's reproductive behavior responds to his environment. This working hypothesis may seem unexceptional to most economists, but many influential voices in the current debate on population policy do not share this view with economists. For example, Professor Ehrlich at Stanford and his associates appear to believe that fertility is *unrelated* to economic forces, and thus, they acknowledge only one avenue for effective policy to deal with high fertility and rapid population growth—namely, replacing individual “free” choice with a system of direct compulsory reproductive quotas.

Economists have contributed relatively little to understanding the causes or consequences of the population problem, and they have not clarified the welfare implications of policies proposed to cope with rapid population growth. It is with a great deal of satisfaction that I note Professor Cain has undertaken this task today, and that Professor Easterlin has called upon his extensive knowledge of U.S. historical data to confirm the environmental responsiveness of fertility and thereby show the feasibility of another avenue for policy—namely, indirect population policies operating through decentralized tax and expenditure schemes.

More specifically, five hypotheses can be distinguished in Easterlin's paper dealing with how differences in environment motivate man to reproduce at different rates. In reverse order they are, first, the knowledge and accessibility of birth control which makes it less difficult or costly to avert “unwanted” births. The second and third hypotheses relate to the net costs of children; it is suggested that food and housing costs of children are greater and the returns from child labor are less in urban than in rural areas. Similar differences in costs and returns are thought to prevail between old and newly settled regions of the country. Two final factors are identified as influencing “tastes,” but it must be noted that Easterlin uses the word taste in a special sense. Educational attainment is alleged to “widen” parent consumption alternatives that are compet-

itive with children, shifting “tastes” from larger to smaller sized families. It is also argued that the larger and more cosmopolitan markets sustained in the settled and urban areas of the country augment the range of consumer goods which again shift “tastes” further toward goods and activities that are competitive with children.

Easterlin's treatment of these two mechanisms under the rubric of changing tastes does not correspond with the economic concept of tastes; changes in tastes or indifference maps are inferred only when changes in relative prices and income do not account for systematic changes in demand. In the first instance, Gary Becker's theory of choice based on the allocation of time provides a persuasive rationale for the commonly observed inverse relationship between a woman's educational attainment and her fertility, given the assumption that children are a relatively time-intensive activity. The market-size hypothesis is more obviously a case where demand responds to change in the relative scarcity prices of goods and activities that are competitive with children. I do not want to discount the potential importance of noneconomic factors that may systematically mold preferences for child bearing and rearing versus alternative activities, but to infer that social, psychological, religious, or genetic factors independently influence parent tastes for children, analysis of non-experimental data requires that the more tangible pecuniary returns of children linked to price and income effects also be isolated and taken into account simultaneously.

Finally, Professor Easterlin comments on immigration and internal redistribution of population as a factor influencing secular trends and interregional differences in U.S. fertility, but a clear understanding of this complex topic will presumably have to await Easterlin's complete discussion of this subject.

Thus, five hypotheses are proposed to account for differences in fertility over time and across seven census regions of the United States. To support these hypotheses the following evidence is noted: (1) fertility, indirectly measured by the ratio of children age 0

to 4 to women age 20 to 44, declined by more than half in 150 years; (2) rural populations generally had higher fertility than urban populations; and (3) frontier and newly settled communities exhibited higher fertility than established and older regions. This evidence may be consistent with one or more of the five hypotheses, but unfortunately no empirical evidence is presented on any of the five determining factors by region or over time.

I suspect that Easterlin's judgment is sound, and that the five environmental factors he notes would help to explain the variation in U.S. fertility over time and across regions; nonetheless, intuition must be verified and much more work is needed. Moreover, quite different conceptual, empirical and statistical approaches than are alluded to here are required to integrate and advance current understanding of the responsiveness of fertility to man's environment.

As I observed with regard to the link between a mother's education and her demand for children, there are elements of micro consumer and producer demand theory which permit one to infer with substantial precision how environmental constraints and opportunities are translated into parent demand for offspring. Also, the process of choice regarding the number of children parents want cannot be viewed in isolation, for this decision is made jointly, to some extent, with a large number of other economic and demographic choices such as marriage, migration, labor force activity, and probably the amount of resources one wants to invest in each of one's children. Analysis of only one behavioral relationship at a time causes both analytical and statistical difficulties; analytically, indirect and feedback effects are ignored which confound causality; statistically, simultaneous-equations estimation techniques are required.

Empirically, more can be done than Easterlin reports in his paper to obtain the appropriate measures of both dependent and exogenous variables. First, the child-woman ratio should be standardized for the age composition of women in different regions and over time. Second, this measure of fertility neglects the fact that parents probably frame their reproductive goal in terms of a number of *surviving* children wanted. Interregional differ-

ences in child death rates appear to still parallel interregional differences in child-woman ratios as late as 1930. According to national life tables constructed for 1900 one-third of the children born 70 years ago would not reach their twenty-fifth birthday, while the proportion today is about four percent. I am curious to know how much of the secular fall and interregional variation in child-woman ratios analyzed by Easterlin are only reflections of the regimes of child mortality parents had to compensate for.

Although Easterlin does not attempt to measure any of his exogenous variables in this paper, series could be constructed in some of his variables. For example, Fishlow's estimates of school enrollment and expenditures by census region for the mid-19th Century do not overtly support the hypothesis that child costs are of dominant importance in understanding differences among child-woman ratios. Clearly the empirical task of constructing long historical series by region is a long and arduous one, but I hope that this effort is begun in Easterlin's forthcoming book.

Finally, a formal statistical model is essential to test the complex multifaceted hypothesis proposed by Easterlin. The test of the hypothesis would then rest on the strength and sign of the partial associations between fertility and each of the properly specified exogenous factors determining fertility. Fortunately a start has been made in this direction in the last few years. A number of econometric investigations have begun to unravel what appear to be economic determinants of fertility, and each has confirmed Easterlin's basic working hypothesis. But as insights accumulate, additional complications also arise—conceptually in modeling the reproductive decision-making process, and empirically in estimating the parameter of the models. The biological and behavioral process of reproduction is likely to continue for some time to challenge economists who are interested in understanding not only the determinants of fertility and population growth, but more generally the economic and demographic behavior of the household sector of the economy.

PAUL DEMENY: Professor Cain's awkward assignment called for presenting a conven-

tional economist's view of the population bomb. Quite sensibly, he cast only the briefest glance at the gruesome spectacle of exploding human masses so dear to the heart of cover designers for popular magazines. When mentioned in the context of a reasoned discussion of population matters, the bomb quickly reveals itself as just another example of vulgar linguistic abuse, hence is easily defused.

Dangerous explosives thus out of his way, Professor Cain could cheerfully proceed towards discharging the second part of his task—to display that he is indeed a card-carrying conventional economist. The terrain he chose for his demonstration—the application of the theory of consumers' choice on decision-making with respect to fertility—is both appealing and useful. Despite the work of Becker, Easterlin, and others, much further elaboration is needed on this theory, and, most of all, much additional empirical work. Professor Cain presents some intriguing fragments from estimates he has developed relating to the economics of childbearing in the contemporary U.S. scene. One wishes that these might be published in detail, hence opened to criticism by economists and demographers. I will resist the temptation to pick bones with some of the assumptions underlying these calculations because my misgivings do not affect the validity of the essential message he wanted to put across—namely that parents weigh, however implicitly, the costs and benefits of having children, make rational choices with respect to their numbers, and, most importantly, that specified changes in the determinants of such choices affect fertility in a predictable direction and to an extent that some day may be quantifiable with reasonable precision. Thus the economist *can* fashion tools, if there is ever a need to use them, to influence fertility. The tools are appropriate shifts of prices and incomes. Preferences or “tastes”—the third factor in the trinity of fertility determinants—presumably could also be a candidate for manipulation, but their murky composition and elusive nature, as well as a number of ideological considerations, makes them, at least for most economists, an unappealing instrument for population policy.

Since even in the social sciences—although, as far as I can tell, not among economists—a

great deal of skepticism remains as to the validity of the economic model of fertility behavior and as to the effectiveness of the policy tools based on it, Cain's summary of the theory in the present context is well justified. Even for economists, often obsessed as they are by their aggregate models, it is a useful reminder that parents' assuming of the so-called “burden” of having children is usually an eminently rational act, since children are a source of utility. But if babies are a manifestation of revealed preference, is there any room left for an interventionist population policy? Aren't those tools for influencing parental choices to be left in their boxes to gather dust forever?

One answer is that there is still much to be done to insure that all children are wanted children, and to insure that certain income and price-affecting arrangements that distort parental decisions on childbearing—such as tax deductions—have been eliminated. Such measures have an obvious first claim on the policymaker's attention. However, measures directed toward these goals represent a population policy only in a qualified sense. They are desirable, and indeed *should* be justified, without any reference to population effects. Reducing costs of birth control or imposing costs of education on parents need not appeal to a wish to reduce population growth.

For the clarity of discussion it is useful then to assume that a state of zero or near-zero cost of birth control has already been achieved and that the consumption of children is no longer subsidized by society. Suppose, furthermore, that under these conditions there results a positive rate of population growth of a substantial magnitude—say one percent per year. I am not saying that such population growth is an inevitable, or even the most likely, outcome of free parental choices. However, Cain's analysis is certainly consistent with the proposition that in an affluent society it is a highly plausible one. Under such circumstances shouldn't then the principle of consumer's sovereignty support an unqualified laissez-faire stance towards individual choices with respect to fertility?

The answer to this question is supplied by observing that rational parents make choices that maximize their *private* utility but that

the consequences of parental choices impinge upon the utility of others as well. In other words, a set of externalities is associated with the consequences of decisions on procreation, an accounting for which does not enter into the calculations of rational parents. Since, save for the most unusual circumstances, these externalities are so diffuse and operate through such extricate mechanisms that they cannot possibly be internalized through private bargaining, there exists a *prima facie* case for social intervention with the fertility decisions of individual parents. Professor Cain does clearly recognize this. Where we part company is in the assessment of the importance of these externalities, hence in the assessment of the strength of the potential case for social intervention.

No one could deny, of course, the enormous difficulties in making such an assessment. Accordingly some judicious hand wringing on the part of Professor Cain would have elicited, at least from this reviewer, much sympathy. At least he could have called for more research on the question, in addition to his pleading for more empirical studies on the income and price elasticities of the demand for children. But we get no expression of misgivings or warnings of caution. Instead, Professor Cain flatly declares that externalities associated with a growing population appear to him "all negative" but "all minor." He does qualify that judgment by limiting its validity to the "near future," but, clearly, he considers that time period the relevant one for policymaking. At any point in time there has always been, and there always will be a "near future" when externalities are labeled minor, thus giving an excuse for the policymaker to wash his hands.

I think Professor Cain is badly mistaken in his judgment on both scores; externalities associated with population growth are not all negative and neither are they all minor. I will not try to argue the case here, but I take it as almost axiomatic that at every time, and in every society, both positive and negative externalities are all-pervasive as well as enormously important. We do not live in isolated families but depend upon, and continually interact—nay, rub shoulders—with our fellow men. The consequences of their reproductive behavior obviously influence our everyday

lives, and our children's, vitally. I cannot accept the proposition that such a crucial element in that behavior as the number of new individuals added to human society, with all the ramifications numbers imply on quality, is a negligible aspect of that influence.

I fear, therefore, that the conventional economist's view just will not do. It seems to me that what the anguished and sometimes hysterical cries about the population bomb reflect is a widespread feeling that externalities are important and, moreover, that negative externalities are now gaining the upper hand; feelings, to put it crudely, that the human population is needlessly multiplying itself at the expense of other species and of nature in general; that aesthetic and qualitative characteristics of our civilization suffer as a consequence of relentless population growth; that the multiplication of human beings maneuvers our society into adopting technical and organizational arrangements that are increasingly brittle and vulnerable to sudden breakdown, and so on. Or, phrasing it in a more positive fashion, feelings that the costs involved in reducing or even eliminating population growth—either through manipulating prices, incomes and tastes, or through some kind of mutually coercive social contract—would be much smaller than the economic, social, aesthetic, etc., gains to be derived from such a demographic change. Whether or not such feelings are correct, they deserve a far more careful examination by economists than they heretofore have received.

It is wholly unsatisfactory to argue that "people" behave rationally in the interest of their children and grandchildren. They obviously do so only as individuals, and there exists no mechanism that would insure that individual decisions somehow add up to a social optimum. To clarify the nature of such an optimum as affected by population, and to devise an efficient machinery that would help achieve it, should be a prime task for economists. Unfortunately, few seem to be willing to take up the challenge. Professor Cain refers to Ardrey's self-seeking rats that end up with a revealed preference for a way of life that he abhors, but expresses the belief that we can choose more wisely than rats. His paper does little to support such a belief.

THE THEORY OF PLANNING

Decentralized Planning Procedures and Centrally Planned Economies

By RICHARD D. PORTES
Princeton University

My title immediately raises difficult questions about definitions. I shall not, however, discuss the meaning of decentralization and the characterization of the extent of decentralization in an economy, subjects which were dealt with at some length in a session of these meetings two years ago (see [9], [15]). We shall have to keep in mind the distinctions between informational decentralization and decentralization of decision-making authority; between the processes of plan construction and plan implementation; between implementing a central plan by prescribing prices (or price-like parameters) or by issuing quantitative commands; etc. But I shall ignore the many unresolved issues in these areas in order to focus on the relevance of decentralized planning procedures to the operation of the centrally planned economies.

By centrally planned economies (CPEs) I mean those of Eastern Europe (excluding Yugoslavia). Here again we must distinguish between the standard Soviet-type system and the variants which have evolved in recent years. Hungary is at present farthest from the standard system, so that the two other papers in this session should indicate the range of institutional frameworks which I have in mind. In this context, I shall point out what seem to me to be the most important problems deserving further research, and then suggest an example illustrating one aspect of a general approach which seems likely to be fruitful in such research.

Decentralized planning procedures are models which assume that the Central Planning Board (CPB) cannot solve the planning problem directly: i.e., either the CPB does not possess all necessary information regarding technology (and possibly preferences), or even if it had this information it could not directly solve the overall optimization problem (because of its size), or both. Hence there must be an exchange of information between the CPB and the various units into which the economy is partitioned, and typically some calculations must be performed by units other than the CPB (for examples of the literature, see [1], [2], [5, Chap. 23], [7], [8], [10, Chap. 24, 25, App. H], [11], [12], [13], [14], [17]). This kind of process does operate in all CPEs, but it bears little resemblance to the form of iterative search for an optimal plan characteristic of the formalized decentralized planning procedures. In practice in the CPEs, plan construction involves very few rounds of iteration between the CPB and lower-level units. Much of the iteration in seeking a plan takes place within the CPB itself, and at the same time there is direct communication between the lower-level units, often going through intermediate levels of an administrative hierarchy.

This kind of apparent discrepancy between the models and the actual planning process raises the broader question of the objectives of the models. In some respects, they are descriptions of planning; in other cases, prescriptions of procedures for

planning; and often, they seem to be primarily computational algorithms for the solution of large-scale programming problems (for a classification and description of most of the better-known procedures from this viewpoint, see [6]). Clearly these objectives are not independent. CPBs do (usually) manage to arrive at plans, so a decentralized planning procedure could not be a very good description of the planning process unless it were computationally tractable. Nor would a planning procedure be a very good technique for planning if it did not take account of the requirements of institutional realism and computational feasibility. Nevertheless, the difficulties in achieving each of these desiderata can be discussed separately. I shall not consider computational questions, except to note that the "programmer's bias"—seeing planning simply as a mechanism for solving a large programming problem—may be responsible for some of the limitations of work in this field.

Those decentralized planning procedures in which prices (or price-like parameters) are the prospective indices sent by the CPB to lower-level units (e.g., [2], [5, Ch. 23]) would seem at first glance particularly unhelpful as descriptions of CPEs. Indeed, one might define a CPE as an economy in which there is central physical allocation of goods in quantitative terms, with prices acting only as tools for aggregation and evaluation of performance. Several recently proposed procedures have therefore used quantitative targets as the indices transmitted by the CPB (e.g., [1], [7], [10], [17]). Yet under the incentive systems conditioning enterprise behavior in CPEs, prices have always played a significant role in influencing the choice of input and output mixes. Of course, these prices are normally fixed for long periods of time and are seldom consciously manipulated by the planners to affect enterprise decisions. But this is equally a criticism of the

quantity-guided procedures, which typically suppose that at each round of the planning process the firms (or sectors) report some variety of shadow prices back to the CPB.

In fact, this is perhaps the most evident disadvantage of existing planning procedures as descriptions of planning in CPEs:—they are too firmly based on duality. The conceptual attractiveness of duality theory has made it the natural foundation for procedures designed to seek an optimal plan. In the CPEs, however, prices are fixed parameters for both the CPB and enterprises, and the information actually exchanged is in the form of quantities, technical coefficients, and financial aggregates.

Duality does imply that with any quantitative allocation proposed by the CPB is associated a set of shadow prices, but in practice these are never made explicit, either by the CPB or by the enterprises. Nor do the enterprises report anything like marginal products, or marginal rates of substitution and transformation, back to the CPB. They do give the CPB information about their technology sets, but in the form of proposed revisions of the quantitative targets or technical coefficients they initially receive. Weitzman's model [17] is closest to practice here, but he must assume that the enterprise counters the CPB's infeasible targets with an *efficient* feasible net supply bundle (and marginal products). This is hardly in the interest of the enterprise, which will instead suggest a plan as far inside its production possibility frontier as it thinks might be credible, so that the CPB will give it an eminently feasible plan on the next round.

The closest one might come to duality in describing an actual CPE would be to view the prices as dual prices in a Leontief model. It is perhaps interesting, therefore, that the most explicit use of iterative algo-

rithms as a computational method has been to employ input-output techniques to calculate average cost prices for comprehensive "price reforms" (e.g., Czechoslovakia and Hungary). It should be noted, however, that foreign trade price relations have been used to adjust the prices derived from these models. This in turn suggests an important descriptive weakness of existing planning procedures—their neglect of foreign trade. This is especially limiting in the context of the Eastern European countries, for all of which (excluding the USSR) foreign trade is of great significance.

The inability to formulate decentralized planning procedures with convincing institutional resemblances to actual CPEs may simply reflect the fact that planning in CPEs has been oriented more towards seeking consistency than optimality or even efficiency. One might think that the growing emphasis on efficiency in the CPEs would make decentralized planning procedures more relevant to these economies. Yet the "economic reforms" in Eastern Europe have definitely not taken the directions indicated by either price-guided or quantity-guided procedures. Rather, they have involved less information exchange with the CPB and more horizontal relations between lower-level units (enterprises or industrial associations), mediated by a newly invigorated market with some degree of price flexibility.

T. Marschak [16] has recently put forward some reasons for this phenomenon: mistrust of the dynamic (convergence) properties of proposed procedures; the magnitude of the informational requirements of even the most "economical" procedures; and the inability to find convincing incentive systems to support proposed procedures. All these are currently among the most important questions being investigated in this field, and considerable progress is being made. The

incentive problem seems the most difficult. It has two aspects, incentives during plan construction and during plan implementation. The individual agent must be motivated both to follow the rules of the procedure and to carry out its part of the final plan. To some extent, these might be separated, and different incentives used for each purpose. But any attempt to do this (or to use the same system for both) must take into account the connection between the two. No agent should have the opportunity to increase the reward accruing to him during plan implementation by distorting the information he supplies during plan construction.

I would add to Marschak's points two considerations which seem to me to be particularly discouraging to those in CPEs who might want to apply these procedures in planning. First, in some procedures the computational burden on the CPB is very heavy indeed. For example, the CPB may have to solve a large nonlinear program repeatedly [11], [17]. The recent literature has tended to emphasize the volume of information transmitted and its costs, but for CPEs with limited computer facilities the computational tasks may be the major problem. Second, theory suggests that under certain technological conditions (approximately constant returns), choices of input and output quantities may be very sensitive to price ratios; and more important, practical experience in Hungary (experiments with multilevel planning) has shown that the shadow prices may be very sensitive to the nature and magnitude of quantitative constraints. This is rather disquieting; it would be very difficult to make choices among investment projects and technological variants in an economy in which relative prices shifted drastically from period to period.

My discussion of decentralized planning procedures as tools for planning in CPEs has so far remained within the context de-

finer by existing models. Common to these models are certain interrelated limitations. They take as given the partitioning of the economy into economic units (decision-making agents). They are explicitly general equilibrium models, and they make no attempt to explore the extent to which some planning decisions might be handled with partial equilibrium methods in a way which would be consistent with events in the rest of the economy. Finally, with few exceptions they treat all sectors (industries, firms) in the same manner. It seems to me that if further work in this area is to yield any results which will be useful to planners in CPEs, or even to those of us who are trying to interpret the processes now unfolding in the CPEs, we must begin to go beyond these limitations.

Investigation of the partitioning of the economy and its effects is important for several reasons. Planners in a CPE do have a great deal of latitude for choice in this regard. Indeed, one of the most difficult problems facing the planners is precisely how to organize the economy. They are relatively free to reshuffle ministries, dissolve or create industrial associations, break up or amalgamate enterprises, and reorganize the lines of authority and communication defining the administrative hierarchy. What criteria should they employ in making these choices? Surely we ought to try to say something about how various kinds of planning procedures would operate in different organizational frameworks.

For example, we have "two-level" and "multilevel" planning procedures, but no analysis of the advantages and disadvantages for the procedures of having more or fewer levels. The number of levels and the number of units at each level will affect the desirable degree of aggregation as information passes up from one level to the next. In turn, the degree of aggregation and other variables depending on the

partitioning will determine the extent to which information is distorted as it passes through the system, and this will affect different procedures in different ways.

Conversely, many of the difficulties which limit the generality and applicability of various decentralized planning procedures might be eliminated or alleviated by changes in the partitioning. The nature and impact of problems posed by the technology will depend on the partitioning: externalities can be internalized; the sum of nonconvex supply sets may be convex. The tasks of calculation within each unit will vary directly with its size, while the tasks of the CPB will decrease if there are fewer units. With what partitioning will the total computational burden be minimized? How might the answer change if a price-guided procedure were substituted for a quantity-guided procedure, or if the rules governing the adjustment of prices or quantities were altered? More generally, decentralized planning procedures have dealt with the exchange of information and instructions between the levels of a given planning hierarchy, with little attention to the possibility of direct contacts between units on the same level. Of particular interest would be efforts to find subsets of units within which certain information could be exchanged, thereby obviating the need for this information to pass up to the CPB.

The work of Bessiere and Sautter [3], [4] is of some relevance here. In considering the separability of a particular optimization problem from an overall planning problem, they were led to investigate how the structure of the economy might be altered so as to facilitate separability. What constraints might be imposed on a given unit, or how might the scope of its optimization problem be extended, in order to make the decentralized solution of that problem compatible with the optimal plan for the economy? They were not dealing

with iterative planning procedures of the kind we are discussing, but their analysis does offer some promising methods for bringing considerations of structure and partitioning into this context. In particular, it suggests that in some cases it may be advisable to combine price and quantity indices, a topic to which we shall return.

A closely related question is how to integrate partial equilibrium problems into the general equilibrium framework characteristic of decentralized planning procedures. It is perhaps in this respect that the literature is most removed from the needs and concerns of the practical planner. In particular, investment project evaluation, choices between technological alternatives, and long-term foreign trade choices typically present themselves as isolated decisions. Yet it is clear that the shadow prices used in making these decisions should bear some relation to those dual to the optimal current allocation. Moreover, in an open economy both must be affected by world market prices, and here again we confront the absence of foreign trade from the models.

Finally, there seems no reason whatsoever why one should expect the appropriate mode of interaction between the CPB and a sector (industry, firm) and the rules governing their behavior to be identical for all sectors. Yet a given decentralized planning procedure will apply the same decision rules and the same type of message sequence to every sector. This is neither sensible nor realistic. Why should a sector exhibiting increasing returns receive and transmit the same kinds of messages as one in which decreasing returns prevail? A similar question arises if we contrast a sector generating significant externalities with one which is technologically independent of other sectors (in regard to externalities, not input-output relationships). In one case, prices

might be the right indices, in another quantitative targets might be preferable. In some cases, the two might be combined, with the sector's use of some inputs (say) regulated by quantitative constraints, while for others it is told only the prices at which they may be purchased. Again, sectors entering heavily into foreign trade might be treated differently from others which do not; or producers of tradeable goods with uniform, well-defined world market prices might exchange different sorts of messages with the CPB from those applying to producers of tradeables made to specifications.

Returning to the nature of the partitioning, the CPB might want to use different sorts of indices in dealing with different levels or in relations between them. Similarly for incentives—need the incentive rules which a ministry should apply to its subordinate industrial associations be identical to those which the industrial association should apply to its subordinate enterprises?

In general, it would seem that we ought to seek decentralized planning procedures which differentiate between units in ways appropriate to their particular characteristics, especially in regard to returns to scale, externalities, foreign trade dependence, and the degree of interdependence (in an input-output sense) between the unit in question and the rest of the economy. There is good reason to believe that this would significantly improve the performance of the procedures. Existing work suggests that departures from neoclassical assumptions on the technology require either the transmission of extra information or the imposition of various constraints on agents. Surely this sort of expedient should be limited, if possible, to those agents which generate the problems.

As an example of this kind of differentiated approach, I would suggest the

case in which increasing returns appear in some sectors but not in others. We know that even if the resulting technology set for the economy as a whole is not convex, there will be appropriate dual prices (reflecting marginal rates of substitution and transformation) associated with the optimal allocation, provided that the optimum is not a corner solution. Loosely speaking, this will be so if the (social) indifference surfaces have stronger curvature than the transformation surface (supposing the latter convex to the origin). We shall assume this to be the case.

We take a model with one resource and two goods. It should be clear how the argument would be extended to more general cases. The first good is produced under increasing returns, the second under decreasing returns. Thus we have

$$(1) \quad \begin{aligned} y_1 &= f_1(x_1), & y_2 &= f_2(x_2), \\ f_1' &> 0, & f_2' &< 0. \end{aligned}$$

We shall contrast a price-adjustment process with a mixed price-quantity index system. It is clear that the former will be unstable, but we shall specify both processes in order to draw the comparison between them.

In the first planning procedure, the prospective indices issued by the CPB are prices of the resource and the goods, and the proposals by producers are demands for the resource (x_1 and x_2) and supplies of the two outputs (y_1 and y_2). Total supply of the resource is given at ω , so excess demand is

$$(2) \quad E_o(x) = x_1 + x_2 - \omega.$$

Prices of the resource and the output are denoted by p_o, p_1, p_2 . The response rule governing both producers' behavior is to set marginal profitability equal to zero, implying

$$(3) \quad p_i f_i' = p_o, \quad i = 1, 2.$$

The CPB is assumed to have preferences regarding the two goods representable by a utility function $u(y_1, y_2)$. For simplicity, we take

$$(4) \quad u(y_1, y_2) = \log y_1 + \log y_2.$$

Taking z_1 and z_2 as the CPB's demands for the goods when prices are p_1 and p_2 , (4) gives

$$(5) \quad z_i = \frac{1}{p_i}, \quad i = 1, 2.$$

Excess demands are then given by

$$(6) \quad E_i(y_i, p_i) = z_i - y_i, \quad i = 1, 2.$$

Ignoring nonnegativity constraints, the adjustment rules governing price setting by the CPB are assumed to be simply

$$(7) \quad \begin{aligned} \dot{p}_o &= E_o(x) = x_1 + x_2 - \omega \\ \dot{p}_i &= E_i(y_i, p_i) = \frac{1}{p_i} - y_i, \quad i = 1, 2. \end{aligned}$$

Now (1) and (3) imply

$$(8) \quad \begin{aligned} x_i &= g_i\left(\frac{p_i}{p_o}\right), \quad i = 1, 2; \\ g_i' &< 0, \quad g_i' > 0 \end{aligned}$$

The time path of the system is then given by

$$(9) \quad \begin{aligned} \dot{p}_o &= g_1\left(\frac{p_1}{p_o}\right) + g_2\left(\frac{p_2}{p_o}\right) - \omega \\ \dot{p}_i &= \frac{1}{p_i} - f_i\left[g_i\left(\frac{p_i}{p_o}\right)\right], \quad i = 1, 2. \end{aligned}$$

It is straightforward to show that in a neighborhood of equilibrium, this is in general unstable, i.e., we do not have convergence.

The alternative procedure uses price indices in the decreasing returns sector and quantity indices in the increasing returns sector. Thus the prospective index issued to producer 1 is y_1 , and the producer reports x_1 and f_1' to the CPB; while

the prospective indices for producer 2 are p_o and p_2 , and the producer reports x_2 and y_2 to the CPB.

For sector 2, nothing is changed. We still have

$$(10) \quad \begin{aligned} p_2 f'_2 &= p_o; \quad x_2 = g_2 \left(\frac{p_2}{p_o} \right), \quad g'_2 > 0; \\ \dot{p}_2 &= \frac{1}{p_2} - f_2 \left[g_2 \left(\frac{p_2}{p_o} \right) \right]. \end{aligned}$$

For sector 1, however, we may write simply

$$(11) \quad x_1 = h_1(y_1), \quad h'_1 = \frac{1}{f'_1}, \quad h''_1 < 0$$

The adjustment rule for the price of the resource then becomes

$$(12) \quad \dot{p}_o = E_o(x) = h_1(y_1) + g_2 \left(\frac{p_2}{p_o} \right) - \omega.$$

We shall suppose that the CPB adjusts sector 1's output target in response to the discrepancy between its own marginal rate of substitution between the two goods and the current marginal rate of transformation. Thus

$$(13) \quad \dot{y}_1 = \frac{\partial u / \partial y_1}{\partial u / \partial y_2} - \frac{f'_2}{f'_1},$$

where the partial derivatives are evaluated at the current output levels. Then (1), (4), and (10)–(13) give

$$(14) \quad \begin{aligned} \dot{p}_o &= h_1(y_1) + g_2 \left(\frac{p_2}{p_o} \right) - \omega \\ \dot{p}_2 &= \frac{1}{p_2} - f_2 \left[g_2 \left(\frac{p_2}{p_o} \right) \right] \\ \dot{y}_1 &= \frac{1}{y_1} f_2 \left[g_2 \left(\frac{p_2}{p_o} \right) \right] - \frac{p_o}{p_2} h'_1(y_1) \\ g'_2 &> 0, \quad h''_1 < 0. \end{aligned}$$

This system does converge to an equilibrium which is a local maximum of u .

We have dealt with the difficulty pre-

sented by the environment in a natural way, using the price mechanism in the circumstances to which it is suited and quantitative targets where prices fail. The procedure has its defects, of course. It is not monotonic in u , nor does it yield a feasible plan if truncated. But this is also true of the pure price-adjustment process. The point here is simply that the specific problem created by increasing returns in *one* sector has been eliminated without having to use quantitative targets for *both*.

This is just an example of one approach to one of the many problems which I have raised. Some of these problems are more tractable than others. If we can make some progress with a few, however, perhaps we may be able to come closer to bridging the gap now existing between decentralized planning procedures and the realities of the CPEs.

REFERENCES

1. M. Aoki, "The Problem of Incentives in the Theory of Planning," paper presented to the Second World Congress of the Econometric Society, Cambridge, England, September 1970.
2. K. Arrow and L. Hurwicz, "Decentralization and Computation in Resource Allocation," in R. Pfouts (ed.), *Essays in Economics and Econometrics in Honor of H. Hotelling*.
3. F. Bessière, "The Concept of Separability and the Optimization of Economic Organization," *Eur. Econ. Rev.*, Fall 1969, 74-91.
4. F. Bessière and E. Sautter, "Optimization and Suboptimization: The Method of Extended Models in the Nonlinear Case," *Mgt. Sci.*, Sept. 1968, 1-11.
5. G. Dantzig, *Linear Programming and Extensions*.
6. A. Geoffrion, "Elements of Large-Scale Mathematical Programming," *Mgt. Sci.*, July 1970, 652-691.
7. G. Heal, "Planning without Prices," *Rev. Econ. Stud.*, July 1969, 347-362.

8. G. Heal, "Planning, Prices and Increasing Returns," forthcoming in *Rev. Econ. Stud.*
9. L. Hurwicz, "On the Concept and Possibility of Informational Decentralization," *Amer. Econ. Rev. Papers and Proceedings*, May 1969, 513-524.
10. J. Kornai, *Mathematical Planning of Structural Decisions*.
11. E. Malinvaud, "Decentralized Procedures for Planning," in E. Malinvaud and M. Bacharach (eds.), *Activity Analysis in the Theory of Growth and Planning*.
12. E. Malinvaud, "Notes sur l'étude des procédures de planification," *Can. Jour. Econ.*, Feb. 1968, 16-36.
13. E. Malinvaud, "Planning for Individual and Collective Consumption," paper presented to the Second World Congress of the Econometric Society, Cambridge, England, September 1970.
14. S. Marglin, "Information in Price and Command Systems of Planning," in J. Margolis and H. Guitton (eds.), *Public Economics*.
15. T. Marschak, "On the Comparison of Centralized and Decentralized Economies," *Amer. Econ. Rev. Papers and Proceedings*, May 1969, 525-532.
16. T. Marschak, "Decentralizing the Command Economy," paper presented to the Conference on Economic Reform in Eastern Europe, University of Michigan, November 1970.
17. M. Weitzman, "Iterative Multilevel Planning with Production Targets," *Econometrica*, Jan. 1970, 50-65.

The Hungarian Economic Reform, Past and Future

By TAMAS NAGY

Institute of Economics, Hungarian Academy of Sciences, Budapest

On January 1, 1968 Hungary launched an economic reform. Its aim was to create a type of socialist economic system in which the planned central control of the national economy was organically combined with the functioning of the self-regulating market mechanism. The basic features of this reform are described in detail in *Reform of the Economic Mechanism in Hungary* (ed. István Friss), Budapest, Akademiai Kiadó, 1969. Brief and comprehensive appraisals were published also by American economists, as by Bela Balassa in *Economica*, February 1970 and by Richard Portes in *The American Economic Review*, May 1970. Because of this I can focus attention here on the achievements and on some essential problems of further development of the Hungarian reform.

The almost three years which have gone by are not sufficient for a detailed and exact evaluation of the Hungarian experiences, but there can be no doubt that our new socialist economic mechanism has proven its superiority over the old one.

In 1970 Hungarian industry is expected to increase its output by 7-8% in contrast with 5% in 1968 and 3% in 1969. An increase as rapid as this sometimes occurred under the conditions of the old mechanism. But at that time it was caused primarily by a large-scale increase in the industrial labor force and led to unavoidable difficulties in realization of the commodities produced, in supply of the population, and in foreign trade. Unsalable inventories bulged and this damaged the efficiency of national economic accumulation. The supply of foodstuffs did not

keep pace with the increased purchasing power of the population, while a suitable commodity composition was also lacking and this led to inflationary pressure. Material imports from Western countries (dollar accounting relation) increased, while increases in exports could not keep pace with this, so the tension of the balance of payments in dollar relation increased, while a possibly too big surplus developed in the ruble relation. In 1970, however, the rapid increase in industrial production is taking place on a more healthy basis. Sales for final use (for domestic consumption, investments in fixed capital and exports) are increasing faster than industrial production. Employment in industry is growing only about 1%. As a result of the increasing output in industry and other branches of the national economy, the national income is expected to be about 6% higher than it was last year despite the poorer agriculture year.

It looks as if we have succeeded in overcoming the chronic problem of exaggerated increase in inventories. While prior to the reform the yearly increase of inventories came to 6% of the GNP, and this was also true of the first year of the reform (1968) as well, this proportion dropped to 3.4% in 1969 and will be about the same in 1970.

But the favorable effects of the new mechanism are developing most rapidly in the field of foreign trade. This is of tremendous importance from the point of view of Hungary, since the weight of foreign trade in our economy is big, the volume of exports coming to over one-third of

the national income. Foreign trade increased rapidly (by 13%) already in 1969, in such a way that exports grew faster than imports, and so our foreign trade balance closed with a substantial surplus even in the dollar relation. The rapid expansion of exports continued in 1970, but now imports are increasing faster (by about 20–22%) than exports, a fact which for this year is not in contradiction to the intentions of the government.

Total money incomes of the population will increase by about 9% in 1970, and wage incomes by about 7%. Inflation has been avoided. In the almost three years which have gone by, the annual increase in producers' prices amounted to about 1%, that in consumers' prices to 1–1.5%.

In comparison with the past these are very good results. However the experience of almost three years also showed that the reform has not as yet been able to cure all the old ailments of the national economy and that the new mechanism itself has certain shortcomings.

The Price System

At the beginning of the reform the sphere of *free* prices was set very carefully, but we hoped to extend their proportion in two or three years. However in this we have been disappointed. Ending certain inherited imbalances in the economy, weakening the monopolistic positions of enterprises, strengthening the pressure of competition and developing buyers' markets progressed at a slower rate than had been expected. Under these circumstances direct governmental price regulations could not have been reduced essentially. Also at present free consumers' prices extend to only 24% of the total retail trade turnover. Taking into account that the state regulation of producers' prices as well as consumers' prices imposes direct constraints on industry, the sphere of free price determination for in-

dustrial enterprises extends to no more than 40% of total industrial output.

The *producers'* prices established by the new price system are undoubtedly better than the old ones. The unrealistically low price level of imported production goods has been ended. The low price level of agriculture has been increased. Relative prices are expressing the differences in capital intensity of production and have become nearer to equilibrium prices.

But our producers' price system remained fundamentally autarkic in character. The prices are mainly based on costs of domestic production and the deviations between domestic and foreign trade prices are "bridged over" by customs duties and export subsidies. So relative prices and differences in profit rates do not give appropriate orientation for development of production and do not exert sufficient pressure on enterprises. Obviously the approach of domestic prices to foreign trade prices can and must take place gradually. This is a process under which liberalization of foreign trade and its prices put pressure on the restructuring of domestic production—and restructuring of domestic production, speeded up by centrally planned and controlled development, makes the gradual liberalization of foreign trade possible. This process would be facilitated by the development of the market-integration of the CMEA countries, in consequence of which the relative prices of this regional market would approach the world market price ratios.

The relative *consumers'* prices for goods and services have not been changed substantially by the reform. They strongly deviate from the producers' prices (and also from the foreign trade prices) by means of differentiated turnover taxes and price supplements. A certain part of these deviations expresses the actual preferences of the socialist state, but many of them are economically unjustified. Besides

this, the general price level of workers' total consumption and consequently the level of nominal wages is low compared to the price level of instruments of labor. This is made possible by the artificially low price level of foods, services and housing. A rise in rents will take place as early as 1971. But a fundamental change in the consumers' price system is an extremely difficult job, particularly in a socialist country where care must be taken to see that a change like this shall not disturb the regular rise in real incomes, and shall not lower—even temporarily—the living standard of a single important social stratum.

The Place of the Firm, the System of Incentives and Income Regulation

The state-owned firms, or so called "state enterprises," supplied 93% of the total industrial output in 1969. They enjoy a wide economic autonomy, do not receive obligatory plan indicators at all, and are operating under market conditions. The state controls and influences their operations mainly by indirect methods—by prescription of some rules for the enterprises, by fiscal policy, credit policy, price policy, etc.

In the highly centralized Hungarian industry, competition is still weak, and for the time being we have more of a buyers' than a sellers' market. But competition is developing. The rapid increase of imports in 1970 strengthened import competition. Many of the enterprises are enlarging the sphere of their activity and entering thereby into competition with each other. This process is not objected to by the branch ministries. In many fields industrial cooperatives, which have developed extremely rapidly in recent times, are playing an increasing role in the development of competition. On the basis of governmental decision, preparations are under way for the establishment of new

state-owned *medium-sized* factories. Finally, in recent times the ministries have begun to dissolve some *large-scale* industrial enterprises, the existence of which proved to be economically unjustified.

In connection with competition and with the place of the firm in our economic system, it is also necessary to draw attention to some important state decisions which already have taken place concerning the forms of capital reallocation. The new legal and financial measures will ease the association of enterprises for certain common activities and the creation of new common enterprises as well. They will permit enterprises to transfer capital to each other from their development funds for investment purposes, either in final form (in which case the enterprise supplying the capital can share in the profits of the recipient enterprise), or in the form of credit (which would have to be repaid together with interest). They will permit commercial credit, prohibited until now—that is, enterprises may grant each other commodity credits or payments in advance from their development funds. Also foreseen in a limited sphere is the experimental issuing of bonds primarily for meeting local (communal) needs.

The present system of incentives and income regulation (which connects the increase in personal incomes of the enterprises' employees with increasing profits) has proved its expediency in some respects. It imposed rational restrictions on the rise of the average wage level and prevented exaggerated income differences between workers in different enterprises which could cause social tensions. But it did not provide sufficient incentives to increase productivity of living labor and it even had a "staff increasing" tendency.

It's true that in 1970 (as mentioned above) the per capita production in industry increased to a satisfactory degree. The "staff increasing" tendency could not be

realized in this year because of the fact that there was no free labor force in the industrialized areas. But tension on the labor market—the lack of labor force (which is, we can say, only a virtual one)—has had an unfavorable effect on work discipline, has increased the harmful shift of the labor force from one job to another, and has hindered its economically advantageous and necessary regrouping.

This is why certain important modifications concerning the regulation of personal incomes will come into force in 1971. Starting next year, enterprises may count the full increase in average wages as costs in calculating their taxable profits, and in accordance with this the rise in average wages will not have to be deducted from the sharing fund. The new rules for the compulsory division of profits enlarge the sharing part but at the same time increase the average tax on this part and decrease the progressivity of the tax. This means that a larger part of the profit increment can be used for increasing personal incomes than formerly. If the enterprise raises average wages it will pay a *special one-time tax* on the raise from the sharing fund. The size of the tax will depend on how the percent increase of average wage compares to the growth of an enterprise index, the sum of annual wages + profit divided by the number of employees. After each percent growth of this index the enterprise can increase the average wage by 0.3%, paying a preferential one-time tax of 50% of the wage increase. Average wage increases running beyond this limit are taxed very progressively: for a further increase up to 0.5% the tax is 150%, for a further increase between 0.5–1% the tax is 200% and so on. These tax rates are not yet definitive. The enterprise may, of course, use the remaining sharing fund to pay bonuses and year-end profit-shares.

The new regulations make it easier for

the enterprise to raise the average wage in an appropriate ratio to the increase of labor efficiency. Along with this, more of the profit increment can be placed in the sharing fund, and so it can be used to increase personal incomes. Even if these new rules do not completely eliminate the unfavorable effects of the old system, calculations have shown that they will lessen the drawbacks without changing the essence of the system of income regulation.

The incentive system for the managers (directors and deputy directors) has already been changed in 1970. Their basic salary was raised through central measures, but now their profit shares are in the same ratio to their basic salaries as those of the workers. Besides this, managers were granted so-called profit bonuses to be paid also from the sharing fund. Their size depends on how the part of profit getting into the sharing fund compares to the annual sum of wages paid out. The bonus is calculated according to a formula but can be modified by supervisory organs depending on how they evaluate the work of the manager. I think this is principally correct. It is not purposeful to connect the rewards of managers *mechanically* to the *annual* formation of profits or other “indicators.” The owners (or their representatives) shall have the right to evaluate the manager’s work and reward him. I do not see any necessity why this job must be done any worse by the branch ministries in socialism than by the board of directors in capitalism.

We can find some economists in Hungary who oppose the basic principles of our income regulation system. They disapprove of worker’s wages (personal incomes) being dependent on enterprise economic achievements (growth in profits or in productivity of living labor), because in their opinion this has an inflationary effect and distorts economic calculations. They want all enterprises to pay

identical wages for identical individual work and to realize this by a mechanism operating in modern capitalism—the collision of interests of trade unions endeavouring to raise wages and of managers striving to increase or maintain profits. This concept has been rejected by the official standpoint. I too consider it incorrect. In a socialist society wage control in principle cannot be built on the collision of interests between workers and managers. Such a mechanism would also be inoperable in practice—the managers couldn't resist pressure for wage increases. On the other hand the fundamental identity of interest between workers and managers can operate as a new and important driving force in economic development. This requires workers' wages (personal incomes) to depend in some extent on the economic achievements of their enterprise. On this basis the actual participation of workers in management can gradually develop under socialism. This is one of the important conditions for diminishing alienation of labor in modern industrial society.

The Investment System

At present total investments of the Hungarian economy amount to about 25% of GNP. One-half of these are state investments (specifically decided by the state) and the other half enterprise investments (decided by the enterprises and cooperatives). In 1969 approximately 48% of total investments were financed from the state budget, 41% from the resources of enterprises and cooperatives and 11% from bank credits.

This too goes to show that the reform brought essential changes also in the investment system. But efficiency of investments did not improve satisfactorily and the system of investments remained a weak point of our economic mechanism. It is a very complex task to achieve an es-

sential improvement in this field of economic activity. A large majority of Hungarian economists is convinced that, besides improving the methods of central planning, it is also necessary to reduce further the sphere of specific state decisions in the investment field, to apply appropriate budgetary methods of influencing enterprise investment decisions and to extend the role of bank credits in financing investments. This indispensable process is hindered, for the time being, by the defects of our relative prices, by the shortcomings of the income regulation system, by the large investment commitments of the state and the investment pressures of the branch ministries. Therefore development in this field has to be slow. For the period 1971–75 (the fourth five-year plan) state investments are planned to amount to one-half of total investments (owing to the growing proportion of non-productive investments, including housing) and to two-fifths of all industry investments (owing to the great share of state investments in projects already started and still in process). In the same period the proportion of bank credits in financing industrial investments will probably increase from the present 13% to about 15–16%. The centralized character of the bank system and its strong control by the state will not be changed, although some modifications of organizational character can be assumed.

At present, as has happened so often in the past, there is a great tension, an excessive demand on the market for capital goods, or more precisely on the building market. In coming years it will be absolutely necessary to achieve a substantial change in this sphere, both by increasing supply (through the more rapid development of the building and building materials industry, through better organization and stronger material incentives) and by

limiting demand (demand of the budget and of the enterprises).

The System of Foreign Trade

As mentioned in my introductory remarks, the economic reform produced a balanced and rapid growth in the foreign trade of Hungary. Also rentability of dollar and ruble exports improved, for enterprises changed the product composition of their exports advantageously. But it is clear that we must go much farther in the differentiation of export expansion, which concentrates resources of development on branches and enterprises with comparative advantages. We must increase the export share of such finished industrial products which can be sold for profitable prices not only on socialist but also on Western markets.

Therefore in 1971 the system of export refunds differentiated for individual *enterprises* will be replaced by refunds differentiated for *branches*. On the other side, limiting imports through administrative measures has shrunk to a narrow field. The import of materials is essentially free also from the dollar relation.

For importing consumer goods from dollar relation, the ministry of domestic trade receives a lump sum of foreign exchange. The compulsory foreign deposits for machinery imported from the West will be reduced to zero in 1971. Also the custom-duty system will be reexamined from the point of view of providing a more favorable effect on the import structure, on technological development, and on increasing competition in the domestic market. And also there is a further intention of this reexamination—to give a better possibility to join the GATT.

Summary

As a summary I can risk the statement that the Hungarian economic reform has passed its examinations well. But stopping the reform at this point would necessarily lead to its regression. Our economic reform must be further developed through carefully conceived measures. The Party and the Government are determined to do this and they can count on the support of both professionals and the vast majority of the population for their endeavors.

DISCUSSION

EDWARD AMES: Twelve years ago, I delivered a paper at these meetings which had probably less impact on the profession than any other paper delivered before or since. It discussed the question of general equilibrium in Soviet-type economies, at a time when it was thought possible to discuss such economies without the use of any theory. It is a great pleasure to note the change that has taken place in the last few years, and, in particular, to comment on Professor Portes' paper.

Portes attempts to relate a new and growing theoretical literature on the theory of economic systems to a literature on the planned economies of Eastern Europe. There are at last grounds for hope that a *rapprochement* between theory and fact may take place in this corner of economics. Portes' paper makes it clearer what problems remain to be solved before the task is complete.

In recent years, theorists have developed a number of models of economics which are not competitive price systems. These include the greed and quasi-competitive processes (Hurwicz), the D-process (Camacho), and the stochastic adjustment processes (Hurwicz-Radner-Reiter). These are the first artificial economies designed by professional economists; while less flamboyant than earlier utopian writings, they are at least intelligible. Moreover, they are analytically interesting, and have certain features which make them attractive to students of Eastern Europe. Portes makes a useful start in exploring their applicability as models of empirical planning systems.

The theoretical literature started from Hayek's conjecture that informationally decentralized economies were more efficient than centralized economies. Hurwicz's first papers attempted to explore decentralized economies which were not competitive price economies. Camacho has recently challenged Hayek's conjecture, for he has constructed a number of very similar systems, one of which is most efficient if completely centralized, a second if partially centralized, and a third if completely decentralized. Since there are no theorems on the conditions in which centralized (or decen-

tralized) systems are efficient, we can use the theoretical literature to organize the discussion of empirical questions, and to knock down the guesses of overzealous empiricists. We cannot, however, consider the subject of centralization as being clearly understood.

Camacho's paper, however, supports the conjecture that the general framework used by Portes is versatile and rugged. In this framework, one examines the flow of information between a *center* (Portes' CPB) and a *periphery* (Portes' enterprises). This flow will be said to converge when the messages in both directions, after sufficient iteration, say (essentially) the same thing, so that center and periphery are agreed on a course of action. This process resembles actual planning practices, and if the iterative process can be shown to converge, then a similar process, truncated after a few iterations, should produce messages which are approximations to the limit message.

An attractive feature of the general model is that Walrasian general equilibrium is a special case of it in which the "auctioneer" plays the role of the center. In this sense, the new literature is a true generalization of economic theory. But there is one important reason for us to be careful about too rapid an application of the model to planned economies.

The literature on competitive prices starts by defining and demonstrating the existence of an equilibrium price, and the stability of this equilibrium. It then shows that a tatonnement mechanism will converge to this equilibrium. In a planned economy, an equilibrium plan is one which (if adopted) would not be changed unless some exogenous part of the economy changes. Presumably an equilibrium plan would be stable if small deviations of plan from the equilibrium would tend to be corrected. A difficulty about discussions of iterative planning is that they ignore the possibility that the iterations will converge to a plan which is not an equilibrium plan. We do not know enough about equilibrium to investigate this possibility, even in simple cases like those studied by Portes.

The practical importance of this question

may be made clear by an analogy. All empirical work on market economies relies on the assumption that observed prices and quantities are approximations to equilibrium values. Empirical work would otherwise be meaningless. Similarly, empirical work on planned economies must ultimately rely on the assertion that observed plans and quantities are approximately equilibrium plans and quantities. But we cannot now assert that this is the case, and the analytical foundations for interpreting the course of events (i.e., for testing hypotheses) are therefore missing in empirical work on Eastern Europe.

Students of Eastern Europe have been interested in the *incentive problem*, which the theoretical literature has not studied. This problem is not entirely new to theoretical economics. In the Walrasian literature, the center (auctioneer) has a simple task—he tabulates excess demands, and uses this information to construct a “next” price vector. The periphery (consumers) in this system have arbitrary (nonidentical) utility functions. In a more recent theoretical development, team theory, the center and periphery have identical utility functions. An incentive problem exists in a team-like system when center and periphery have different utility functions. Many practical difficulties in Eastern Europe seem to exist because of incentive problems. It is for this reason that Portes speaks of the importance to planners of adjusting the utility functions of enterprises so as to make the economies more like teams. But it would be interesting to know more precisely what can be the effects of such adjustments on the operations of an economy, and to know that, we need an extension of team theory.

A third problem discussed by Portes arises from another observation by East European specialists. The empirical East European economy consists of a variety of organizations having a hierarchical structure of parts. Montias has recently written an interesting paper on this subject. To consider such an economy in terms of existing theory, one breaks the hierarchy into two parts, its upper end being the center, and its lower parts the periphery. It may be arbitrary at which point this break is made (just as in systems analysis there may

be a few big black boxes or a lot of little black boxes).

Now let us ask whether it matters how the hierarchy of the center is organized. If the plan is genuinely a general equilibrium system, then the answer should be *no*. But it is widely held that empirical plans are in fact compilations of partial equilibrium plans; and in this case it may well matter how the center is structured. Thus when, in 1957, the USSR replaced national industrial ministries by regional economic units, many economists conjectured this change would affect economic life. The new theoretical literature treats the center as an indivisible atom, and existing models, in Portes’ words, “are explicitly general equilibrium models, and they make no attempt to explore the extent to which some planning decisions might be handled with partial equilibrium methods. . . .” Ultimately this question must be explored, and it will involve modification of existing models.

These three problems arise from attempts to reconcile experience with theoretical constructs. They point the way to further development of comparative economic systems. They would not have been formulated unless some interaction had taken place between theoretical and empirical economists. This interaction has taken place in a part of the discipline in which no theory at all existed until quite recently. Portes’ paper is therefore a most encouraging event in the growth of economics.

BELA BALASSA: Tamás Nagy has given a balanced account of the Hungarian experience with the economic reform since January 1968. I am in general agreement with his conclusions and I too consider the results of the reform to have been basically favorable. I will not comment therefore on particular points in his presentation; rather, I will attempt to reformulate the problems Nagy has raised in terms of a simple model and will indicate the future policy changes necessary for the full implementation of the model.

The model underlying much of official and unofficial writings on the reform prior to its introduction could be characterized as one of a socialist market economy where government

intervention is largely limited to the use of indirect policy measures, such as fiscal and monetary instruments. It was assumed that competition among socialist firms would ensure that profit maximization by these firms would be in the national interest.

This model differs from the familiar Lange model in two important respects. First, instead of price setting by a central authority, price formation was assumed to be increasingly left to market forces. Second, in contradistinction to Lange's closed economy model, the Hungarian model is "open." Under the small country assumption of international trade theory, it was envisaged that Hungary would specialize in accordance with world market price relations.

In a closed economy, decision making by individual economic units will lead to a Pareto optimum if all units are price takers, there are no external economies, and factors of production can move freely among firms. Welfare economists have been much concerned by the lack of fulfillment of the first assumption due to the existence of economies of scale on the firm level. But increasing returns to the firm will not create a problem in a small open economy whose firms are price takers on the world market. Thus, assuming that adjustments are made for external economies and resources are free to move, profit maximization by individual firms would ensure a Pareto optimum in the Hungarian model. At the same time, in an open economy with efficient-size firms there is no conflict between the requirements of static efficiency and technical progress as would be the case in a closed economy where the atomistic competition, necessary for ensuring that all firms are price takers, would hardly be conducive to technical progress.

The practical application of this model in Hungary has been subject to a variety of constraints, however. Firstly, the government wished to avoid wage and price inflation. Second, there was a desire to avoid creating large income inequalities and—to rephrase Nagy's statement somewhat—lowering the living standards of any social group, however small. Thirdly, the government intended to minimize the uncertainty associated with rapid structural changes. Last but not least, social orga-

nizations such as local party and trade union units as well as vested interests in the supervising ministries and in the individual firms circumscribed the freedom of action at various levels of decision making.

The existence of these constraints largely explains the imposition of what are called "brakes" in Hungarian economic writings. They include the maintenance of price fixing over a large group of commodities, the widespread use of differential subsidies and taxes in foreign trade, the application of regulations penalizing wage increases, the limitations imposed on bonus schemes for management, the virtual maintenance of the existing highly concentrated industrial structure, the limitations on the movement of capital among firms, and the continuing operation of every existing firm. Despite the progress made, these limitations have interfered with the efficient operation of the market mechanism in Hungary so that profit maximization by individual firms does not necessarily correspond to the requirements of social profitability.

This situation has given rise to conflicting recommendations. Some have suggested that since, in such circumstances profits do not provide an appropriate indicator of the firm's performance, additional success criteria need to be applied by the supervising ministries. By contrast, others have taken the view that one would have to establish the conditions for the identity of interests on the firm and on the national economy level so as to conform to the model of the socialist market economy.

The first alternative has little to commend it. Decision making by the firm responding to additional criteria, necessarily based on a partial view of the economic process, is bound to create inefficiencies in resource allocation. Moreover, firms would again spend much effort in bargaining with the supervising ministry to reduce and to circumvent the prescribed targets. Some signs of this are already apparent in bargaining for export subsidies, tax exemptions, and bonuses.

Correspondingly, intervention by the ministries should be reduced rather than increased. Moreover—and here I seem to differ with Nagy—I would consider it desirable to separate the state's regulatory function from its

function as the owner of capital. This could be accomplished by replacing the branch ministries with a single industrial ministry that would have a general regulatory function and transferring the supervisory-ownership functions to a board of directors on which representatives of the state, the management, and the workers would be represented. Such boards should be given power to discontinue the operation of the firm, to establish joint subsidiaries, and to substantially increase operations by the use of borrowed capital.

As regards the regulatory function of the state, much remains to be done to increase domestic competition by breaking up multiplant firms whose continuing existence is not warranted by scale economies. Moreover, as Nagy correctly points out, there is need to increase import competition and to ensure that price relations on the world market are translated into domestic prices without being distorted by differential subsidies and price-support schemes.

The attainment of the latter objective encounters difficulties, however, by reason of the duality of Hungarian foreign trade. Broadly speaking, Hungary trades in two markets—designated as ruble trade and dollar trade—where relative prices differ to a considerable extent. As Alan Brown and Paul Maser have shown in a recent paper, prices of materials tend to be higher in the West whereas machinery prices are higher in the East. If Hungary faced infinitely elastic foreign demand and supply curves, the appropriate solution would be to follow the old rule: buy cheap and sell dear and provide incentives to firms accordingly.

The assumption of infinite foreign elasticities certainly holds as far as dollar imports are concerned and applies more-or-less also to dollar exports. By contrast, the availability of materials in ruble trade is limited through bilateral agreements; moreover, proceeds from

exports to socialist countries are not convertible.

It is necessary therefore to devise a system of pricing goods and foreign exchange that will maximize Hungary's gain from trade while maintaining balance-of-payments equilibrium in dollar trade as well as in ruble trade with the individual socialist countries. While the solution to this problem would require a general equilibrium framework, some simple rules of thumb may be suggested. First, the single rate of exchange (in Hungarian parlance, foreign exchange conversion ratio) applied in ruble trade should be supplemented by a system of import taxes and export subsidies on trade with the individual socialist countries so as to equilibrate Hungary's balance of payments with each country. Second, a system of taxes and subsidies should be applied to trade in individual commodities which are traded with both the East and the West so as to ensure that the prices firms pay for imports equal the marginal cost of these imports to the economy and the prices received for exports equal marginal revenue. In practice this would mostly mean that firms would be confronted with the prices applicable in dollar trade.

Similar considerations have led the Hungarian Planning Bureau to introduce the use of dollar prices in evaluating major investment projects. And while the proposed system of taxes and subsidies may appear complicated, its use will be desirable as long as quantitative and balance-of-payments constraints apply in ruble trade and this is carried out at prices different from those applying in dollar trade. The need for such a system will disappear as the proceeds from exporting to socialist countries become convertible and if and when, as Nagy notes, "through the development of the market integration of CMEA countries . . . relative prices in this regional market would approach the world market price ratios."

THE ECONOMICS OF POLITICAL DECENTRALIZATION

Towards a Predictive Theory of Intergovernmental Grants*

By DAVID F. BRADFORD and WALLACE E. OATES
Princeton University

Until quite recently, theoretical analyses of the impact of intergovernmental grants on public expenditures have run either implicitly or explicitly in terms of the familiar theory of individual choice.¹ The recipient of the grant is typically viewed in effect as an individual decision maker with preference patterns of the usual sort defined over private and public goods. Within this theoretical structure, a number of propositions have been developed including, for example, the conclusion that, dollar-for-dollar, a matching grant will induce a greater expansion in spending on the public good than will a lump-sum, unconditional grant.

The difficulty with this analysis, as we all have been well aware, is that this is not the appropriate theoretical framework for the study of intergovernmental grants. Such grants are not grants to individuals; they are grants to groups of people. This means that the effects of these grants depend upon the process by which the group makes collective decisions. A real theory of intergovernmental grants must, for this reason, be one which takes explicitly into account the political process through which the collectivity determines its levels of spending upon public goods.

* The authors are grateful for a number of helpful comments to the participants in economics seminars at McMaster University and at Rutgers University. In addition, they are indebted to the Ford Foundation whose support has greatly facilitated this study.

¹ For an excellent presentation using this approach, see James Wilde.

In one recent paper, Goetz and McKnew have shown for a special case that it is conceivable that lump-sum grants can have a greater stimulative impact on public spending than an equal-dollar matching grant. Using a rather complicated model in which the collectivity decides separately on the aggregate level of public spending and on the mix of public programs, they show that, if individual preferences bear a particular kind of relation to one another, simple majority voting will lead to greater public expenditure in response to a lump-sum grant than to a matching grant of the same amount.

Whether this special case is likely to occur with any frequency is another matter, but, at any rate, it is clear that one can concoct particular instances in which a process of collective decision making will lead to results which are at variance with the conclusions which follow from the model of individual choice. Our purpose in this paper is to try to explore somewhat more systematically the effects of these two types of intergovernmental grants.

The question arises whether there are interesting classes of collective decision processes for which the comparative effects of the two types of intergovernmental grants can be predicted. In an earlier paper on revenue-sharing (Bradford and Oates), we have constructed a framework for conceptualizing collective budget determination which offers the possibility of dealing systematically with the effects of

intergovernmental grants under varying sets of assumptions about the nature of political processes and the range of political possibilities. In that paper, we examined the case of lump-sum grants to collectivities, and we succeeded in distinguishing a class of political processes under which such grants are equivalent (both in terms of their allocative and distributive effects) to a set of grants directly to the individual members of the collectivity. In this paper, we explore the application of this same approach to a comparison of the effects of lump-sum and matching grants to political units.

In section I, we provide a brief recapitulation of the theory of intergovernmental grants in terms of the standard model of individual choice. This, as it turns out, proves to be of some value for the case of collective choice as well. Section II describes the conceptual framework for analyzing processes of collective choice that we developed in our earlier paper and indicates the relevance of the basic theorem proved there to the problem under study here. In section III, we use this framework to analyze the allocative and distributive effects of lump-sum and of matching intergovernmental grants where the collectivity makes its fiscal decisions by simple majority rule. In particular, we prove that, under simple majority rule with fixed tax shares and a single public good, a matching grant will always lead to a larger public expenditure than will a lump-sum grant of the same amount. In conclusion in section IV, we speculate on the prospects for generalizing this result.

I. Intergovernmental Grants in a Model of Individual Choice

In this approach, the recipient of the grant is typically assumed to be an individual decision maker with convex preferences defined over quantities of a private

good and a public good as in Figure 1.² To simplify the analysis we assume that units of both goods are chosen such that the price of the public good in terms of the private good is unity.

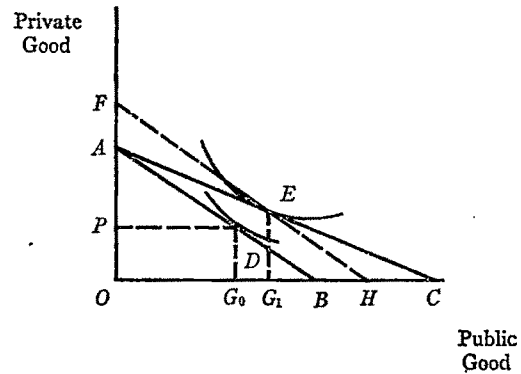


FIGURE 1

Subject to a pre-grant budget constraint of AB , the community selects an output of the public good of OG_0 which leaves OP to be consumed in the form of the private good. Assume next that the community is the recipient of a matching-grant program; the central government, for example, might agree to share the cost of providing the local public good by funding a specified percentage of the community's expenditure on the good. In terms of Figure 1, such a grant pivots the community's budget constraint outward about the point A to reflect what is now in effect a lower unit price of the public good to the local community. As a result, provision of the public good in Figure 1 rises to OG_1 , and the community receives a grant of DE .

Suppose, however, that instead of a matching grant of DE , the grantor chose to give the community a lump-sum grant of the same amount. Such a grant would

² Alternatively, one might assume that Figure 1 represents community indifference curves. The difficulties associated with such curves are such, however, as to render them of dubious value, particularly for predicting the effects of government grants.

shift the community's budget constraint outward to FH but would not alter the relative prices of the private and public goods. Moreover, it is clear from Figure 1 that the matching grant would induce a smaller increase in spending on the public good than would the matching grant; since FH passes through E but is steeper than AC , it follows, given normal convexity properties of the preference map, that the tangency of FH with an indifference curve must lie to the left of E . Or, in economic terms, the lump-sum grant induces more expenditure on the public good because it not only enriches the community (i.e., has an income effect), but in addition reduces the price of the public good to the recipient (i.e., has a substitution effect). In the individual model, a matching grant thus possesses a greater stimulating effect on public spending than does a lump-sum grant of the same amount. We now turn to a reconsideration of the properties of these grants in a framework with explicit collective decision making.

II. *The Analytic Structure*

For simplicity, we will at this point continue to use a model in which there is a single pure public good, a single private good, and a given and constant price of each where the units of these goods are defined such that the price of the public good in terms of the private good is unity. We need next to describe in conceptual terms the local community and its political process. Briefly, we represent what we will call the *state of the community* by a vector $[y_1, y_2, \dots, y_n, g]$, whose first n components indicate the disposable incomes (or claims to the private good) of each of the n members of the community and whose last component, g , is the number of units of the public good. The set of *feasible states* of the collectivity represents the community's budget constraint; for each total quantity of the private good available for distribu-

tion among the members of the community, the boundary of this set indicates the maximum level of provision of the public good attainable.

Not all feasible states for the community, however, will be political equilibria. A state, for example, in which all the community's resources were used for provision of the public good would probably generate political pressures which would lead to another state in which members of the community would consume a positive quantity of the private good. In this framework, we can view the *political process* (or the rule for collective decision making) as the mechanism through which one feasible state is transformed into another (possibly itself). For a given set of feasible states, a political process thus defines for each initial feasible state the resulting equilibrium state; it is a mapping which indicates how the political process transforms each state of the community into an equilibrium state. Note in particular that, starting from a given feasible state, the political process simultaneously determines both the level of output of the public good and the disposable income of each member of the community.

From this perspective, we can characterize a grant to the community as a disturbance to the existing equilibrium state and a change in the feasible set. We could think of the grant as initially taking the form of an increment to the level of provision of the public good. However, this new state, which is itself a member of an enlarged set of feasible states, will not in general represent a political equilibrium. It will itself typically be transformed by the political process into a new equilibrium state.

We have found it extremely useful for analytic purposes to explore one particular question relating to the response of the community to a grant program. Specifi-

cally, we ask—in the case of each grant to the community as a whole—if *there is some particular way the grantor might have distributed the grant funds directly to the individual members of the community so that precisely the same equilibrium state of the community would result.* If this is the case, we can show that a particular inter-governmental grant is precisely equivalent, both in terms of its ultimate impact on the provision of the public good and on the disposable income of each member of the community, to a set of grants made directly to the individuals themselves.

We find (not surprisingly) that the answer to this question depends upon the properties of the political process. In our earlier paper, we developed a proposition for lump-sum grants which may be described briefly as follows. Assume that a grant program to a community acts in the first instance to increase some collective activity (where “holding a cash balance in the public treasury” counts as a collective activity). In contrast, assume that each of the set of grant programs to individuals which duplicates the effective set of possibilities (i.e., the feasible set) generated by the grant to the collectivity, has the initial effect of simply increasing individual disposable incomes, leaving the levels of the local government activities unchanged.³ Then, our proposition states that, under a set of suitable conditions on the political process (conditions discussed at length in our earlier paper), there is a program of grants to individuals which leads via the political process to precisely the same equilibrium state of the community as does the grant to the collectivity.⁴

³ We could be comparing, for example, a grant of \$1 million to a local government which initially simply increases local public revenues to a set of grants directly to the individual members of the community where the sum of these latter grants is \$1 million.

⁴ It is interesting that the proof of this proposition does not require that there exist any particular relationship between individual preferences and collective choice.

In the most general case, it is not possible to derive the precise character of this program of individual grants. However, in particular cases in which the political process is specified in some detail (as we will show in the next section), this can be done. Finally, and directly to the point of our study here, we have found that for certain cases, we can link individual preferences and the implicit set of individual grants through the political process in a manner which allows us to predict the effects of changes in the form of the grant on the collective choice. And it is to this issue we turn now.

III. *Intergovernmental Grants and Simple Majority Rule*

An important example in which we can employ our analytic framework to examine in some detail the effects of different programs of intergovernmental grants is the case where the community makes its collective fiscal decision by simple majority rule. More specifically, we will take as our model a community in which each individual's tax *share* is known and fixed and in which the level of provision of the public good is determined by a simple majority voting rule. By this we mean that the community votes on alternative pairs of provision of the public good until one is found which at least $((n/2)+1)$ members of the community prefer to any other. This model could represent, for example, a local school district which finances its expenditures through a local property tax and which employs majority voting to determine the annual school budget. In this instance, each individual's tax share is defined by the assessed value of his property divided by the total assessed value of taxable property within the district. His tax bill is then determined by the product of his tax share and the size of the school budget selected by the electorate.

In Figure 2, we indicate the position of

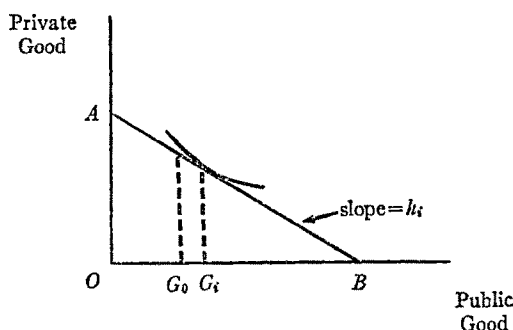


FIGURE 2

any individual (say the i^{th}) member of the community. Since by construction the price per unit of the public good in terms of the private good is unity, the slope of his budget constraint, AB , is equal to his tax share, h_i (i.e., the "tax-price" to him of a unit of the public good). Assuming that each person's preference map exhibits indifference curves of the usual shape, it follows that the preferences of each member of the community for the public good will be single-peaked. In terms of Figure 2 where OG_i is the individual's most preferred level of provision of the public good, this means that, between any two levels of the public good greater than OG_i , he prefers the smaller of the two (i.e., the one closer to OG_i). Similarly, for any two budgets less than OG_i , he prefers the larger. With preferences which satisfy this property of single-peakedness, the theorem of Duncan Black states that, under simple majority rule, the equilibrium budget is the median peak, or in other words, the median of the most-preferred levels of provision of the public good. This means that, aside from the individual who possesses the median peak, the equilibrium public budget will not be that budget which is most preferred by any given member of the community. In Figure 2, for example, the equilibrium budget, say OG_0 , will in general differ from the individual's most preferred budget, OG_i .

Let us now disturb the political equi-

librium represented in Figure 2 by the introduction of a grant program to the community. We consider first the case of a lump-sum grant to the collectivity. Assume that the central government provides the community with an unconditional grant of a specified sum. In Figure 3, if the grant were $EH = BD$, this would imply that, if each member of the community were to maintain his level of consumption of the private good at its previous equilibrium level, the output of the public good could increase to OG_1 , or by the full amount of the grant. The full range, how-

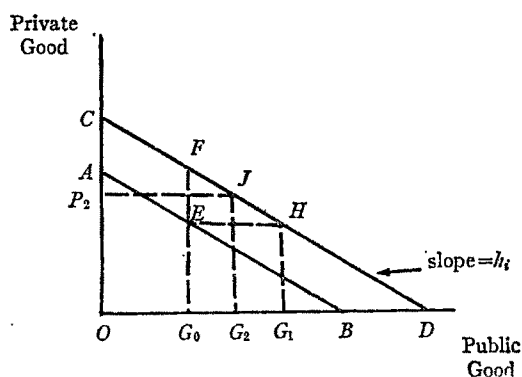


FIGURE 3

ever, of alternative bundles of private and public goods available to the i^{th} individual is indicated by the new budget constraint, CD , and it is unlikely that the point H will represent a new political equilibrium. Suppose instead that, using majority rule, the community decides to increase the public budget to OG_2 so as to increase also consumption of the private good. In Figure 3, point J by assumption thus indicates the new equilibrium provision of the public good and the level of private consumption, OP_2 , of individual i .

We now ask if, instead of using an inter-governmental grant, the grantor could have generated the identical equilibrium state of the community by parcelling out the grant funds directly to the individual members of the community. The answer

is yes. If, for example, the central government had given person i in Figure 3 a lump-sum grant of EF , this individual would have had precisely the same budget constraint, CD , as in the case of the grant of EH to the collectivity. With the same budget constraint and a given preference map, his array of preferences for public budgets and hence his voting pattern would be exactly the same in the two cases.

Note next that $(EF) = h_i(EH)$, the grant to each individual, which from his position is equivalent to the grant to the collectivity, is equal to the product of his tax share and the intergovernmental grant. Moreover, since the tax shares (the h_i) sum to one, it follows that the sum of the grants to the individuals equals the amount of the original grant to the collectivity. This means that, if the central government simply divides up the grant monies among the individuals in proportion to each person's local tax share, precisely the same state of the community results as if the total of the grant funds were given to the public treasury of the community. A lump-sum grant to a community in this model is, therefore, an implicit set of grants to the members of the community where each individual's grant is proportional to his local tax share.

Consider next the case of a matching grant to the community. In Figure 4 with AB again possessing slope h_i and representing the individual's pre-grant budget constraint and with G_0 the original equilibrium public budget, the grant pivots the individual's budget line about point A to AC . Note that the slope of AC is $[h_i(1-m)]$, where m is the fraction of the unit cost of the public good funded by the grantor. If, for example, the grantor adopted a 1:1 matching grant, this would imply that $m = \frac{1}{2}$, or that the effective tax-price to the individual is now:

$$h_i(1-m) = h_i(1 - \frac{1}{2}) = \frac{1}{2}h_i.$$

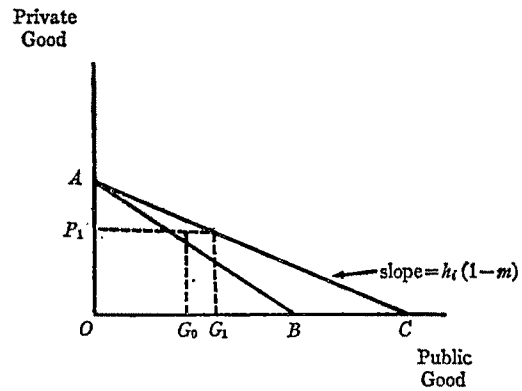


FIGURE 4

A matching grant to the community thus reduces the tax-price per unit to each taxpayer by the fraction m . As a result of the grant, the community will, by majority rule, select a new level of provision of the public good which we indicate by G_1 in Figure 4. In this particular case, our i^{th} person ends up consuming OG_1 of the public good and OP_1 of the private good.

It is easy in this case to see how the central government could have generated this same equilibrium state of the community by dealing directly with the individual members rather than by a matching intergovernmental grant. In particular, suppose that the central government refunded directly to each individual the fraction m of his local tax payment. Or, perhaps more realistically, the central government might allow the i^{th} individual a credit of (mT_i) against his tax bill from the central government, where T_i is the i^{th} person's local tax bill. In either case, the effect on each individual in the community would be identical to that resulting from an intergovernmental matching grant where the central government funded the fraction m of the local budget. In all these cases, the resulting budget constraint for each individual and hence his voting pattern would be the same. It is thus clear that the median peak of preferred budgets would be OG_1 in Figure 4

applicable in this instance; a lump-sum grant to him of EF (or, equivalently, an intergovernmental lump-sum grant of EK) would lead him to a most-preferred provision of the public good which is smaller than that under the matching-grant program.

Consider next, however, the case in Figure 6 where an individual's most-preferred budget under the matching grant is indicated by point M , a budget which is less than that of the median voter. In this instance, it is possible that this person

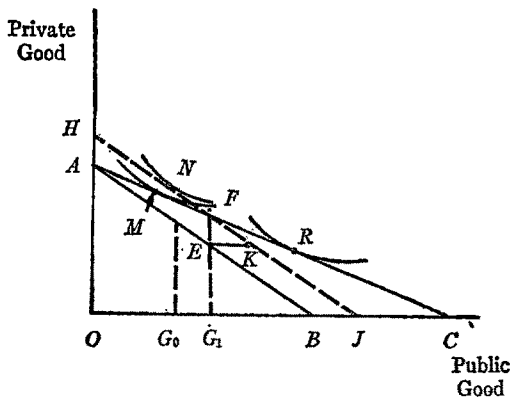


FIGURE 6

could prefer a larger budget (e.g., that indicated by point N) if the community received a lump-sum grant of EK than under the matching-grant program. This can occur because the individual's implied grant, EF , under the lump-sum intergovernmental grant exceeds the implicit sum he would receive under his most-preferred budget under the matching grant. In this case, the added income effect of the larger implicit lump-sum grant may be stronger than the substitution effect of the *hypothetical* matching grant. Note, however, and this is crucial to the argument, that, although such a person's most-preferred budget may be larger under the lump-sum intergovernmental grant, it will never be as large as the equilibrium budget under the matching grant. It is clear from

the geometry in Figure 6 that the tangency of an indifference curve with a budget line under the lump-sum grant must always occur on a curve which is higher than that which passes through M , and this means that the point N must always lie to the left of F .

We have shown, therefore, that all voters with most-preferred public budgets less than the median under the matching grant will continue to vote for budgets under the lump-sum grant which are smaller than the equilibrium budget under the original matching grant. Since the model of individual choice shows that the median voter under the matching grant will himself prefer a smaller budget under the lump-sum grant, it follows that the median of the most-preferred public budgets will be smaller under a lump-sum intergovernmental grant than with a matching-grant program of the same amount. It is, incidentally, easy to see that people who prefer relatively large budgets under the matching grant (e.g., like point R in Figure 6) will generally prefer smaller budgets under the lump-sum grant. For these individuals, the most-preferred budget under the matching grant implies a larger income effect than the lump-sum grant, as well as the price effect.

IV. Some Further Thoughts

In Section III, we were able to employ our analytic framework to reach some specific conclusions concerning the allocative and distributive effects of intergovernmental matching and lump-sum grants for the case of simple majority rule with fixed tax shares. While we feel that these results are of some interest in themselves, they are obviously quite limited in a number of respects. Besides applying to a single collective decision rule, they are derived in a model which assumes convexity of preferences, a single public good, etc.

This raises the issue of the extent of the generality of these results. We might mention that, although we have decided not to burden the reader with another special case, we have been able to prove that the results derived in section III also follow for the case of the Lindahl budget-determination model under the assumption that the collective activity is a normal good. Aside from this, however, we can report only somewhat uneven progress in our attempts to generalize these propositions concerning lump-sum and matching grants.

In our earlier paper, we were able to demonstrate the basic theorem of the existence of a set of individual lump-sum grants which is equivalent to a lump-sum grant to the collectivity as a whole for a fairly broad class of collective decision processes in which there are any number of collective goods and in which there are virtually no restrictions on the properties of individual preferences. The same formal analysis implies that, for this same broad class of political processes, there exists a set of tax credits to individuals and lump-sum redistributive payments within the community which would duplicate the allocative and distributive effects of any particular intergovernmental matching-grant program.

These propositions concern political choices within a *given* set of alternative states. The principal concern of this paper, on the other hand, is the comparison of choices made under different sets of

feasible states, a comparison of the choice made among the alternatives available under a matching grant with that made from the different set of states possible under a lump-sum grant. The restrictions which we placed upon the political process in our earlier paper seem, we feel, both plausible and not obviously equivalent to the conclusions expressed in the propositions. However, to compare outcomes in different feasible sets requires further assumptions about the political process, and it remains an open question whether there are reasonable and interesting restrictions which will ensure in general the result derived for the particular case in this paper.

REFERENCES

- D. Black, "On the Rationale of Group Decision Making," *J. of Pol. Econ.*, Feb. 1948, 56, 23-34.
- D. Bradford and W. Oates, "The Analysis of Revenue-Sharing in a New Approach to Collective Fiscal Decisions," to appear in a forthcoming issue of the *Quart. J. Econ.*
- C. J. Goetz and C. R. McKnew, Jr., "Paradoxical Results in a Public Choice Model of Alternative Government Grant Forms," in *The Theory of Public Choice: Essays in Application*, J. M. Buchanan and R. S. Tollison, eds. (Ann Arbor: University of Michigan Press, forthcoming).
- J. Wilde, "The Expenditure Effects of Grant-in-Aid Programs," *Nat. Tax J.*, Sept. 1968, 21, 340-8.

State and Local Response to Fiscal Decentralization

By A. JAMES HEINS
University of Illinois

When inquiring into the area implied by the title of this paper, I first ran into the problem of specifying what was meant by decentralization. My first thought was that it was similar to the process undergone in recent years by the armed services and some business firms. In this context decentralization implies that the larger organization confines itself to setting the broader goals of policy, while turning over authority to make specific decisions to ingredient units of the organization under a well-defined set of rewards and punishments. For example, instead of allotting specific types of new equipment and supplies to its ships, the central command in the Navy might allot sums of money, leaving the spending—thus the decision how those resources might be used most effectively to increase the ship's worthiness—to the ship's captain. The rewards and punishments used by the Navy to provide the necessary behavioral stimulus to the ship's officers are fundamental to this process, but need not be detailed here owing to their familiarity. But a moment's reflection leads one to conclude that fiscal decentralization in the U.S. is quite different since it involves sovereign ingredient units. That is, the nature of the rewards and punishments by which the federal government imposes its policy on states and localities is vastly different from that internal to most organizations. To be sure, the federal government can offer carrots in the form of matching grants, but it cannot compel the kind of tax-expenditure performance by states and localities consistent with the usual notion of decentralization.

With those qualifications, I view the process of fiscal decentralization as a process under which decisions about scope of activity are made at the federal level, but decisions about the design of activity are made at lower levels of government. The immediate implication is that fiscal decentralization involves programs under which funds are acquired by the federal government and turned over to lower levels of government for disposition.

(Parenthetically, one might note that this process seems to involve simultaneous centralization—federal taxes rise—and decentralization—funds are turned over to states and localities. In fairness, however, it should also be noted that the first major interest in massive grant programs surrounded the expectation of a post-Vietnam fiscal dividend. Turning over surplus funds to states rather than increasing federal spending would seem to be legitimate decentralization).

Whatever the term used, it is clear that federal deployment of funds to states and localities is an increasingly significant fiscal phenomenon. And, it is also clear that good theory is in short supply. Usually we approach the problem by assuming governments behave as individuals; that is, we employ voters' preference functions or global utility functions with all the attributes of individual functions and think in terms of the familiar income and price effects.¹

¹ Goetz and McKnew [4] note that this approach "is so widely accepted and taught that it is difficult to cite published attempts to formally establish its base." They note as an exception the controversy in [3, 7, 8, 9, 10]. For a more recent example of this approach see [6, pp. 31-54].

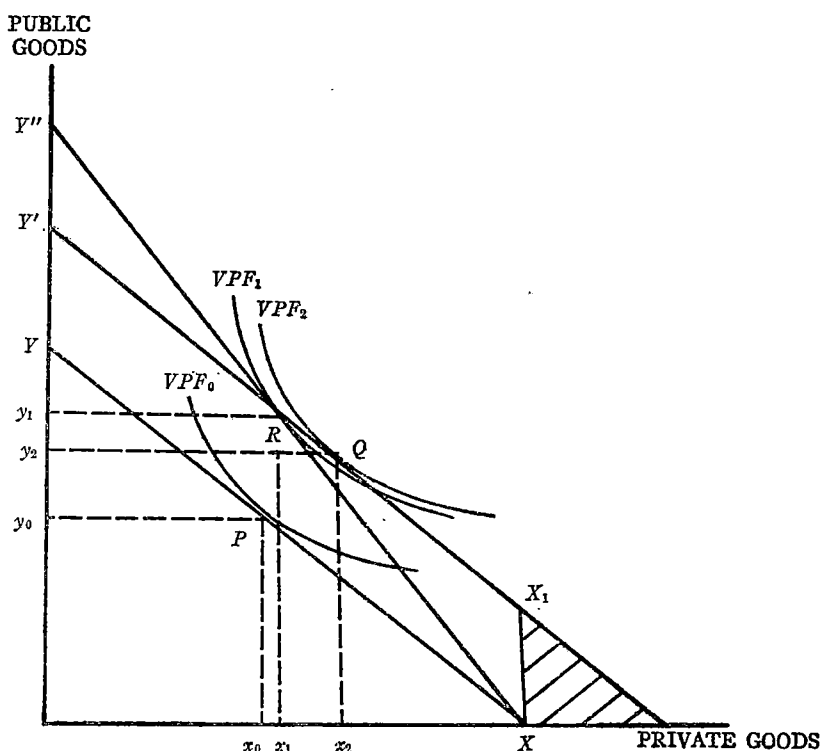


FIGURE 1

The essence of the theory is embodied in Figure 1. Assume some community initially faces budget constraint FX as between public and private goods. Given voters preference function VPF_0 , the community takes position P and divides its resources up such that x_0 private goods are acquired and y_0 public goods. If the community receives an unconditional grant equal to $Y' - Y$, the effect is to shift its budget constraint to $Y'X'$, and it moves to point Q where it acquires x_2 and y_2 of private and public goods respectively. (If the grant cannot be rebated to individuals and spent on private goods, the effect is to stop the budget constraint at point X' and make the shaded area unattainable.) If both private and public goods are normal, then:

$$y_2 > y_0, \quad x_2 > x_0, \quad \text{and} \quad y_2 - y_0 < Y' - Y.$$

These inequalities merely state that the flat grant will partly result in an increase in public goods acquisition and partly in tax relief, the precise results depending on the relative income elasticities of demand for alternative goods.

A matching grant to the community of exactly the same amount will induce a shift in the budget constraint to $Y''X$. The matching terms are usually stated as $(Y'' - Y)/Y''$, the share of expenditures on public goods paid for by the grantor. The terms necessary to make the alternative grants equal in amount are those required to take the community to point R where VPF_1 is tangent to $Y''X$ at a point on budget constraint $Y'X'$. (The fact that these terms might be difficult to determine merely means the amount of money granted under a matching grant depends on community behavior. A flat grant is

precise as to amount.) If the voters' preference functions have the attributes normally assigned to individual indifference curves, then it must be true that: $y_1 > y_2$ and $x_1 < x_2$; or put another way, point R must be northwest of point Q . This analysis merely predicts that a matching grant of a given amount will result in greater expenditures on public goods than a flat grant of the same amount. The reason is that the matching grant has a positive price effect going for it in addition to the income effect.

This is the essence of the controversy in [3, 7, 8, 9, 10]. Note that in [3, p. 103], the authors state that any other result (than $y_1 > y_2$) indicates that: "the jurisdiction would be responding irrationally to the grant offerings." Here the authors are not only claiming that the voters' preference functions look like indifference curves, they indicate that deviations indicate irrationality on the part of the community. In fact, they refer to this community behavioral pattern as representing a utility function of some sort. Their problems in making these inferences I think are transparent.

Goetz and McKnew get at the real problem in [4]. They offer an illustration of a three-person community, each being rational, having well-defined rules for collective choice, in which a flat grant results in greater public expenditures by the community than a matching grant. The real problem is that the state of our knowledge about the process of collective choice does not entitle us to make inferences about the nature of voters' preference functions, and certainly not that they look like indifference curves. What these functions look like depends on the preferences of individuals and the collective choice process. Furthermore, one can only talk meaningfully about individual rationality and not about collective rationality.

Having said all this, I think it is highly

likely that matching grants are more "stimulative" than flat grants. Goetz and McKnew merely demonstrate that this cannot be proved theoretically, just as Giffin demonstrated that one cannot prove that individual demand curves must slope downward. Given this state of the theory, the question of the relative power of various grant forms to stimulate public expenditures remains an empirical question.

I think questions of even greater import lurk behind the problem of predicting the effect of grants on aggregate expenditures. For example, what effect will various grant forms have on tax sharing arrangements at the state and local level?² What effect might grants have on the propensity for rich people to migrate to wealthy tax havens? How might the very state and local decision-making process be affected when the bulk of funds come from outside while expenditures are determined within? These questions arise because grants tend to fracture the nexus between costs and benefits in public decision making.

To illustrate how these problems might be approached, consider Figure 2 in which I depict a community in which each of three people (poor, average, and rich respectively) have demands D_1 , D_2 , and D_3 for public goods (assumed to be pure). Assume that the community is at a stable equilibrium in which all excess demands have been resolved and, thus that at \bar{q} the relative tax shares are t_1 , t_2 , and t_3 such that $t_1 + t_2 + t_3$ equals the marginal cost of another unit of the public goods. Assume also that tax payments $t_1\bar{q}$, $t_2\bar{q}$, and $t_3\bar{q}$ are proportional to the respective taxpayers' incomes.

Suppose now that the federal government raises its progressive income tax and uses the funds to provide a flat grant to the community sufficient to purchase q^* of the

² I raise these questions with respect to predicting the impact of a massive income maintenance program in [5].

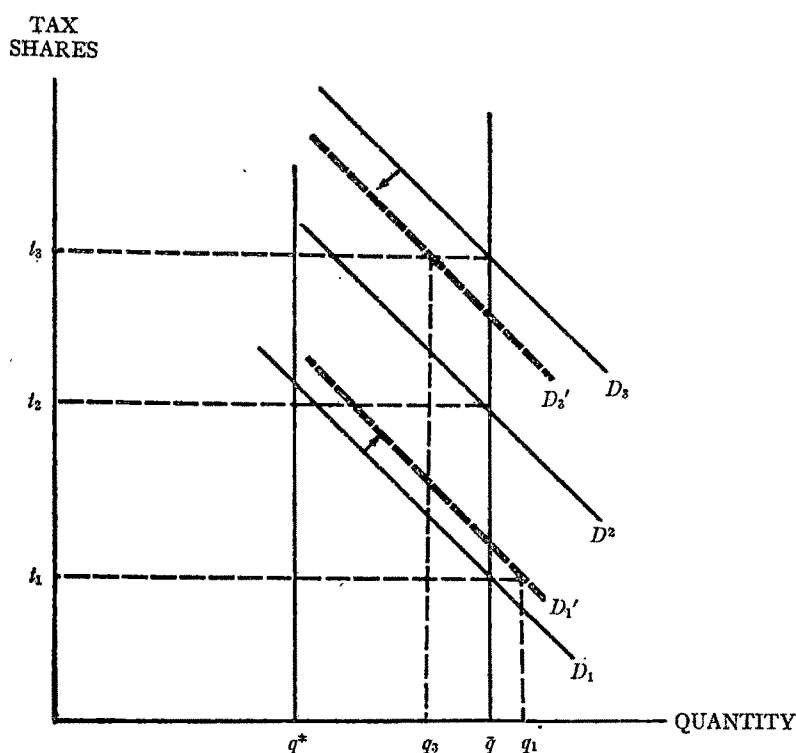


FIGURE 2

public good. That is, the illustrative community receives a grant of $(t_1 + t_2 + t_3)q^*$. Assume also, that this community is a typical community in the sense that the sum of federal income taxes collected from the three residents precisely equals the grant. Now, if f_i represents the incremental effective federal tax rate on individual i and y_i his income; the effect of these assumptions is that:

$f_1 y_1 < t_1 q^*$, $f_2 y_2 \leq t_2 q^*$, and $f_3 y_3 > t_3 q^*$ and:

$$f_1 y_1 + f_2 y_2 + f_3 y_3 = (t_1 + t_2 + t_3)q^*$$

In terms of the demand for community public goods, the income effects tend to reduce rich man three's demand to D_3' , increase poor man one's demand to D_1' , and leave the effect on median man two indeterminate. Here I assume $D_2' = D_2$. The net effect of the illustrated tax-grant

program is to create excess demands equal to $(q_1 - \bar{q})$ and $(q_3 - \bar{q})$ for individuals one and three respectively.

Generalization: *When an individual pays more (less) taxes to the grantor under a fully-funded, flat-grant program than he pays to the grantee for equivalent services, the algebraic value of his excess demand for public service of the grantee will be reduced (increased).*

While the above generalization is straightforward, it is difficult to make further generalizations about its impact on collective decisions of the grantee. This is precisely the new theory we need so badly. But, not despairing, let me say some things that can be said.

For the moment, I will assume tax shares of the grantee are fixed and less progressive than those of the grantor. I also assume that preferences for public goods and the collective choice mechanism are such that

the personal distribution of excess demands can be represented by a continuous, bell-shaped (not necessarily normal or non-skewed) curve with the median point at zero. This is tantamount to assuming that the median voter prevails in the decision-making process. The grant to be considered is a flat grant of a given per capita amount.

Now, residents of wealthy communities will experience predominantly reduced excess demands for public services and poor communities increased excess demands. Assuming an even distribution of excess demands among the rich and poor, the predicted effect would be for the flat grant to increase public spending in poor communities and reduce it in wealthy communities, leaving average communities pretty much unchanged. The same result could be readily derived from the simple model illustrated in Figure 1. The typical community would experience no essential change in its budget constraint; taxes paid out to the federal government would equal the grant received. The budget constraint of wealthy communities would shift back to the origin, and for poor communities the opposite would occur.

The above conclusions would, of course, arouse immediate objection from many economists since they suggest that the full effect of the grant is to create a tax cut by the grantee, with public spending changes resolved through residual income effects. Objections to my first conclusions might be based on the notion of "tax illusion;" that is, once taxes are on the books and people adjust to them, few votes can be purchased with a tax reduction while many votes can be lost with a tax increase. Thus, windfall dollars tend to be spent instead of being rebated to taxpayer. This notion is often supported by such casual empiricism as: "Have you ever seen taxes cut because of a grant-in-aid?" I guess the answer is that I have never seen a grant immediately large

enough to expect a discernible effect. The "tax illusion" notion is based upon assumptions of an irrational and ill-informed electorate. While perhaps correct at the moment, they cannot be a basis upon which to create interesting theories of collective behavior.

A more sophisticated case for the positive impact of a flat grant can probably be made on the basis of interjurisdictional tax competition. Since the effect of a grant system is to pay for more state and local spending out of federal taxes, the inframarginal gains from shifting location because of favorable tax shares are less. For example, in Figure 2 if taxpayer three were to flee this community to another where his tax share would be t_2 (instead of t_3), the tax saving prior to the grant would be $(t_3 - t_2)\bar{q}$. After the grant system is installed, however, his tax saving would only be $(t_3 - t_2)(\bar{q} - q^*)$. The difficulty here is that jurisdictions also have less reason to offer tax concessions since they derive fewer revenues under the grant system for any given reduction in tax share offered. Thus, the aggregate effects are not clear.

I think more important considerations arise, however, when I relax the assumption that excess demands of all kinds are experienced by rich and poor alike. In Figure 1, for example, we assume the grant to merely shift the budget constraint and leave the voters' preference functions unchanged. However, once we invoke distributional considerations and a collective choice process, the latter assumption cannot hold.

To illustrate, consider an average (in terms of income) community in which all the well-off people have excess demands clustering about the zero mark, while the not-so-well-off have excess demands ranging from large negative to large positive. In this case, the effect of the grant may be to shift the distribution of excess demands such that the expenditure level declines.

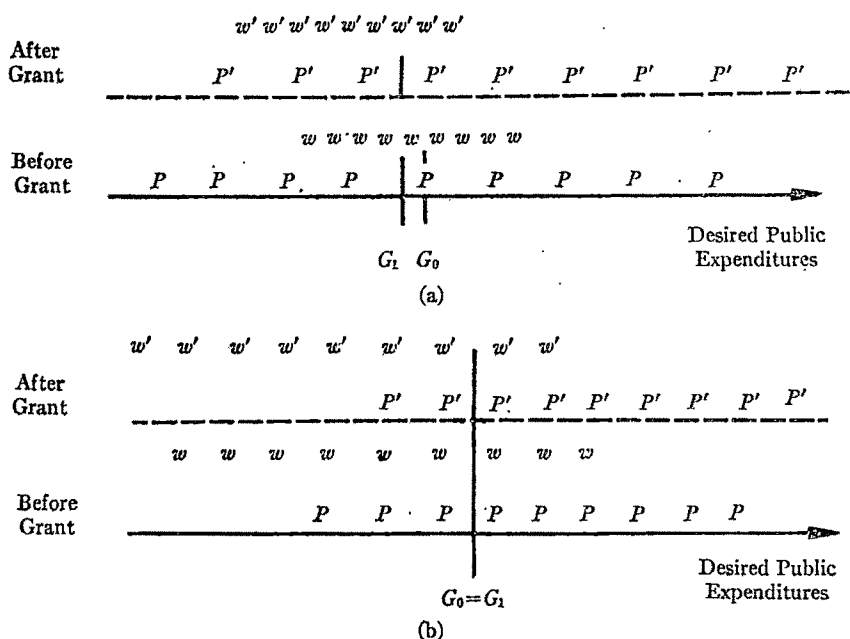


FIGURE 3

Such a situation is illustrated in Figure 3 (a) (before grant case) where the symbol w represents a position of a well-off person on the desired expenditure scale and the symbol p the position of a poor person. This being a typical community there are as many p 's as w 's. Assuming fixed tax shares and dominance of the median voter, the expenditure level will be G_0 ; and all persons represented to the left of G_0 have negative excess demands and all persons to the right of G_0 positive, while those at G_0 have zero excess demand.

Suppose we now adopt a grant system which changes excess demands as indicated in the after-grant case. Here a w' represents the new position of a well-off person to the left of his previous position, the algebraic value of his excess demand, relative to G_0 , having decreased. Similarly p' represents the new position of p (always to the right) after the grant. Here I have assumed that the absolute values of the changes in excess demand are the same for each individual. Note that Figure 3 (a) shows that the expenditure level will de-

crease from G_0 to G_1 . We get this counter-intuitive result despite the assumption that this is an average community and despite the assumption of the similarity of effect of excess demand changes on all people. It happens because the distribution of excess demands among the rich and poor are as indicated. Many social critics might invoke the greater mobility of the rich (a fact which tends to reduce excess demands as shown in [11]) and their greater political power to suggest that Figure 3(a) does not represent an atypical distribution of excess demands. But it should be abundantly clear that the excess demands could have been distributed in Figure 3(a) so as to generate an increase in expenditures as a result of the grant.

Suppose excess demands had been distributed as shown in Figure 3(b). In this case no change in expenditures is indicated; however, the dispersion of excess demands has increased. In other words, a flat grant may not only affect the median level of excess demands, and thus the expenditure

level, but also the dispersion of those excess demands. Some interesting consequences may follow from this effect.

If we imagine a social variable called "political tension" and think of this as measuring the extent to which residents of a community are dissatisfied with political decisions, then we can think of the results illustrated in Figure 3(b) as indicating an increase in political tension. Some manifestations of this result might be: (1) migration of wealthy community members to wealthy tax havens in order to resolve their increased negative excess demands; (2) social conflict in the form of riots, boycotts, etc.; (3) an increase in the regressivity of community taxes as the poor offer up larger tax shares in order to resolve positive excess demands.

Again, it should be made clear that the distribution of excess demands could have been drawn such that the grant would reduce the dispersion of excess demands, and thus lead to results very different from the three posed above. The point is that we do not know enough about the distribution of excess demands for community services to resolve these questions.

Note also that the effects outlined interact. Migration may ease social tension within a community. Changing tax-sharing arrangements may reduce the dispersion of excess demands and thus induce the wealthy to stay put or generally ease social tension. (See [1] for a discussion of the strategic aspects of this consideration).

What this comes down to is that the possibility of major employment of grants-in-aid by the federal government as part of a program of political decentralization poses complex questions that must be seriously considered by social scientists. It is not simply a question of Washington sending funds to the states so they can afford to buy the public services they need but cannot afford. Rather, a major grant program threatens systemic change in political decision making at the state

and local level resulting from the fact that grants obscure the relationship between costs and benefits of public enterprise. It would be nice to know about more these prospects.

REFERENCES

1. James M. Buchanan, "Principles of Urban Fiscal Strategy," Research Paper No. 808231-1-2, Center for Study of Public Choice, Virginia Polytechnic Institute, Blacksburg, Virginia.
2. James M. Buchanan and Gordon Tullock, *The Calculus of Consent* (Ann Arbor, Michigan, 1962).
3. John Cotton and Thomas O'Brien, "Unconditional Versus Matching Grants: Further Comment," *National Tax Journal*, March 1968, pp. 103-104.
4. Charles J. Goetz and Charles R. McKnew, Jr., "Paradoxical Results in a Public Choice Model of Alternative Government Grants," mimeograph, Center for Study of Public Choice, Virginia Polytechnic Institute, Blacksburg, Virginia.
5. A. James Heins, "Income Maintenance and the State and Local Tax-Expenditure Package," Discussion Paper No. 53-69, Institute for Research on Poverty, University of Wisconsin, Madison, Wisconsin.
6. Selma J. Mushkin and John F. Cotton, *Sharing Federal Funds for State and Local Needs* (Praeger: New York, 1969).
7. Alan Nichols, "Unconditional Versus Matching Grants: Comment," *National Tax Journal*, September 1967, pp. 344-345.
8. Lester C. Thurow, "The Theory of Grants-in-Aid," *National Tax Journal*, December 1966, pp. 373-377.
9. ———, "Unconditional Versus Matching Grants: A Reply," *National Tax Journal*, September 1967, p. 346.
10. ———, "A Further Reply," *National Tax Journal*, March 1968, p. 105.
11. Charles Tiebout, "The Pure Theory of Local Expenditure," *Journal of Political Economy*, October 1956, pp. 416-424.

Centralization and Decentralization: The Political Economy of Public School Systems

By O. A. DAVIS and K. O. KORTANEK
Carnegie-Mellon University

The issue of centralization vs. decentralization has a long history in the literature of economics. It was, for example, central to the debate about the superiority of the capitalist or socialist system. Essentially, this issue was a major concern of Von Mises, Hayek, Lange and others. This issue is also central to the arguments of the more recent and more mathematically oriented theorists who have addressed it more directly. See, for example, Marschak [14], Whinston [15], and Hurwicz [11]. In addition, it has been the subject of extensive investigation in the management science literature. See Baumol and Fabian [2], Charnes, Clower and Kortanek [3], Charnes and Kortanek [6], Hass [10], Kortanek [13], and Whinston [16].

In recent years the issue has attained increased urgency and practical importance due to the advocacy by important elements of our population of a decentralized organization of certain of our public school systems. For example, the Mayor's Advisory Panel on Decentralization of the New York City Schools [17], intended that through decentralization the following goals be achieved: (i) increase community awareness and participation in educational policy, (ii) open new channels and incentives to educational excellence, (iii) achieve greater flexibility in the administration of the schools, (iv) give a large city a single school system that combines the advantages of big-city education with the opportunities of the finest small city and suburban educational sys-

tems, (v) strengthen individual schools and encourage close coordination with other private and public efforts to advance the well-being of pupils and all others.

It was proposed [17] that the school districts have authority for all regular elementary and secondary education within their boundaries and responsibility for adhering to state education standards. Each district should be large enough to offer a full range of educational services and yet small enough to promote administrative flexibility and responsiveness to community needs (including their diversity from district to district). A central education agency in coordination with other governments should have operating responsibility for special educational functions and city- (or region-) wide policies. It should also provide certain centralized services to the school districts and others upon the district's request.

Despite the laudatory objectives of some of the advocates of decentralized public school systems, this immediate issue does not appear to have attracted the same kind or level of scholarly attention accorded the more general controversy of centralization or decentralization. This paper represents an attempt, perhaps the first scholarly attempt, to address the abstract issue of decentralization at approximately the level indicated by the example of the controversy over the organization of the public schools. The approach followed here, which is based upon that developed in Charnes, Clower and Kortanek [3],

differs from traditional economic analyses in the sense that the emphasis is not upon market mechanisms of the tatonnement or other variety. A "Central Educational Authority" (abbreviated CEA) is assumed to transmit information to districts (or subdistricts) and the latter commits resources according to an optimization with respect to variables under its control.

Since the chosen tools of analysis are those of mathematical programming, there is an obvious necessity of having an objective function. Clearly, several strategies could be employed in the design and specification of such a function. Perhaps the most obvious approach would be to follow in the tradition of orthodox welfare economics and claim that the objective is to maximize an ethically acceptable social-welfare function which has individual utilities as its arguments. Not only because of the practical difficulties, but also because of our belief that this approach would yield little of interest, it was not pursued here. At the opposite extreme one could attempt to construct a function based upon, say, the stated objectives of the Mayor's Advisory Panel which were summarized above. Due to obvious difficulties, as well as the evident possibility that the objectives may not command wide support, this approach too was not pursued. Instead—and realizing that this strategy may be controversial—we choose to specify several alternative objective functions. All of these alternatives are simple. Each has some intuitive appeal as a practical surrogate for a social welfare function. Presumably, at some other place and time we might argue about which, if any, of these might be appropriate for actual use.

At this stage our interest is the development of a model which will incorporate some (certainly, not all, and perhaps not even the most pertinent) features of this local and somewhat untypical (for eco-

nomics literature) type of decentralization. Using the Eastman and Kortanek [9] school facility model as the basic building block, an attempt is made to relate the analysis to the dynamics of community development over a long-term planning horizon, and to the fundamental interrelatedness of our communities. Attention is centered upon such variables as the rates of local property taxes, their interrelationships, the sequencing of construction of housing units, etc. Perhaps the most fundamental weakness of this attempt is our ignoring the basic issue of the quality of output—educational achievement, whatever it is—by presuming (as did the Mayor's Advisory Panel [17]) that "acceptable" quality levels are specified by a higher authority and enforced via the device of standardized tests. Yet, we argue that a first step is necessary and that the proposed structure is an interesting beginning for the provision of a structure in which this issue of decentralization might be analyzed in full.

Notations for School Population Dynamics of a District

We shall use the Eastman-Kortanek model [9] which is directed towards the dynamics of housing transitions and aging of children in a given school district. That model represents the school population of a given district as the product of three kinds of dynamic processes: (1) building of new homes in the district and their purchase by families, (2) natural transitions in the housing stock, and (3) the birth and aging of children in families living in existing housing within the district. Space constraints require that the reader refer to Eastman-Kortanek [9] for the development and exposition of the basic ingredients of this model.

We assume that there are $\theta = 1, 2, \dots, \bar{\theta}$ school districts and that each is governed by possibly different dynamics of the same

general form. We repeat below some of the important school district decision variables which will be involved in the collection of all districts.

Class Sizes for District θ

In accordance with section 5 of [9],

- (1) $e^T S_\theta^{(t)}$ = the vector of the total number of children of all ages in the community, and the i^{th} component
- (2) $(e^T S_\theta^{(t)})_i$ = total number of children of age i in school district θ .

Therefore the number of children age i added or removed from school district θ is given by

$$(3) \quad (e^T S_\theta^{(t)})_i - (e^T S_\theta^{(t-1)})_i.$$

Actually these variables are dependent on others, and will enter in constraints on class size later on.

Annual Costs of Education

It has been recommended [17] that funds be allocated to a school district in terms of "unit needs." A "unit need" is defined as the demand on the school system by one year of education for one child. It appears possible to develop an information system measuring needs, setting weights among these and determining improvement by perhaps a battery of standardized city-wide skill tests.¹

Our first purpose here however is in developing and studying structure of inter-district dynamics and therefore we shall assume that the degrees of differing pupils needs in districts can be measured by a cost for accommodating one student of age i during time period t . We shall denote this

unit cost for school district θ by:

$$(4) \quad c_{i\theta}^{(t)} = \text{cost of accommodating a student of age } i \text{ during period } t.$$

This measurement reflects some agreement of "capacity," e.g., class size, number of teachers needed for this level, etc. In addition, however, this cost includes extra tutoring or remedial work necessary to bring the pupil to his level of acceptable performance on standard skills tests. We shall assume that these unit costs differ period-by-period, allowing the possibility of their decrease relative to higher start-up type costs usually encountered initially.

Schedules for New Housing Construction

The sequences of new housing construction are variables for each school district. We do not mean to imply that these decisions are to be made by the school board of district θ . In fact the Eastman-Kortanek [9] model was developed to respond to friction that commonly occurs between the developer, homeowners, and the school board of a school district. Nevertheless we let

$$(5) \quad (N_{1\theta}^{(t)}, \dots, N_{k\theta}^{(t)})$$

denote the total number of new houses of type 1, ..., k respectively built in period t in district θ . Our purpose of structure study is to be able to ascertain the effects of a variety of such schedules in respective school districts on the total interrelated collection. Different sequences will attract different families and hence impose various dynamics on the school population over time. The same is true due to the natural dynamics of already existing housing, even if, say, there were little or no new construction.

Budgetary Considerations

Sources of funds to support education are tax revenues from houses within the

¹ An information system developed along these lines but for different purposes was developed by Touche Ross & Co., for the Mayor's Committee on Human Resource Development for the city of Detroit in 1967.

district and various subsidies from governments or the central education agency (CEA).

Let

$$(6) \quad F_t^{(\theta)} = \text{funds given to districts from the CEA.}$$

Let

$$(7) \quad \tau^{(t)T} = (\tau_1^{(t)}, \tau_2^{(t)}, \dots, \tau_k^{(t)})$$

be the vector of yearly market values of each house type in district θ at time period t . The market value of all houses is

$$(8) \quad M_\theta^{(t)} = \tau^{(t)T} [I_k I_k] H_\theta^{(t)},$$

where H^t denotes the $2k \times 1$ vector (2 kinds for each type: sold and not sold) of the total stocks of housing types in district θ at time t , and I_k denotes the $k \times k$ identity matrix (k = number of housing types in each district). See 6.3.1 of [9].

Let

$$(8.1) \quad \mu_\theta = \text{tax rate on assessed value in district } \theta.$$

Then

$$(8.2) \quad \text{Total tax revenue at time } t, R_\theta^{(t)} = \mu_\theta M^{(t)}.$$

We shall impose the following set of inequalities between education costs, tax revenues, and funds subsidies (for each district θ):

$$(9) \quad F_\theta^{(t)} + \sum_{i=1}^n R_\theta^{(t)} \geq \sum_{i=1}^a c_{i\theta}^{(u)} (e^T S_\theta^{(u)})_i, \quad \text{for } u = 1, 2, \dots,$$

$$(9.1) \quad c_\theta^{(t)} = \sum_{i=1}^a c_{i\theta}^{(t)} (e^T S_\theta^{(t)})_i = \text{cost of instruction in unit } \theta.$$

Thus in period u , the total of cumulative

tax receipts plus the subsidy F_θ^u must be bigger than or equal to total cumulative educational expenses.

We shall assume that there is a minimum number of newly constructed houses for the entire complex consisting of all school districts over the entire planning horizon. These demand profiles may be gathered from population studies for different age groups, who usually require varying characteristics in housing attributes depending on age, income, number in family etc. Thus, we shall require

$$(9.2) \quad \sum_i \sum_\theta N_{\theta i}^{(t)} \geq m_i,$$

where m_i is defined to be the minimum number of i -type newly constructed houses required for the total region during the planning horizon.

Various Constraints Including Interdistrict

Class Size Targets Over Districts

For purposes of illustration we develop here constraints on each age group over different districts. Let K_θ^L and K_θ^U represent the total number of students that can be accommodated with present facilities at the "junior" and "senior" levels respectively where ages 1, 2, ..., ℓ comprise junior level L and ages $\ell+1$, $\ell+2$, ..., a comprise senior level U . Since we are interested in allowing shifts on movements of age groups of one district into another should the dynamics so indicate, for example, the combining of previous districts into new ones, we include a capacity constraint over all districts with the intent that the K_θ^L and K_θ^U bounds may be relaxed to a reasonable degree to permit the desired regroupings. Hence we write:

$$(10) \quad \sum_{\theta=1}^{\bar{\theta}} \sum_{i=1}^e (e^T S_\theta^{(t)})_i \leq K_i^L \quad \text{for each } t$$

$$(11) \quad \sum_{\theta=1}^{\bar{\theta}} \sum_{i=l+1}^a (e^T S_{\theta}^{(i)})_i \leq K_l^U \quad \text{for each } l$$

These constraints also have implications on the number of required teachers (vis-a-vis target class sizes).

Overall Budgetary Limitations

In (9) above we defined $F_{\theta}^{(i)}$ as the subsidy in period i to SD_{θ} required to cover the difference between cumulative tax revenues and cumulative school expenditures at time i .

We introduce the following coupling constraints

$$(12) \quad \sum_{\theta} F_{\theta}^{(i)} \leq A_i \quad \text{for each } i$$

and

$$(13) \quad \sum_{\theta, i} F_{\theta}^{(i)} \leq A.$$

Note that (12) is a constraint on total in a given period, while (13) is a constraint over all school districts and communities.

Uniform Tax Rates vs District Tax Rates:

Some Game Theoretic Solution Concepts

We now propose alternative objective functions to guide the study of tradeoffs between a policy of a single tax rate μ across all districts vs possibly differing tax rates for individual districts, μ_{θ} . Some of these objectives are suggested in part by developments in [1] and their price-theoretic interpretations for expenditures of local government.

$$(14) \quad \bar{\mu} = \text{"uniform tax rate"}$$

defined by the State (see also [7], p. 98).

A Uniform Rate

Let μ be the uniform rate which is to apply to all districts. Consider the following optimization:

$$(15) \quad \text{Min } \mu$$

subject to

$$\mu \geq (C_{\theta}^{(i)} - F_{\theta}^{(i)})/M_{\theta}^{(i)},$$

$$C_{\theta}^{(i)} - F_{\theta}^{(i)} \geq 0, \quad \text{for all } i \text{ and all } \theta;$$

$\mu \geq \bar{\mu}$, and all the district and intradistrict constraints of the preceding section.

Thus, (15) seeks the smallest tax rate μ above the uniform tax rate which when applied to the market value in each district will cover total district instruction costs less state subsidy.

Another form of linear fractional type is the following:

$$(16) \quad \text{Min } \mu$$

subject to

$$\mu \geq \left(\sum_{\theta} \sum_i [C_{\theta}^{(i)} - F_{\theta}^{(i)}] \right) / \sum_{\theta} \sum_i M_{\theta}^{(i)},$$

$$C_{\theta}^{(i)} - F_{\theta}^{(i)} \geq 0 \quad \text{for all } \theta \text{ and } i,$$

$$\mu \geq \bar{\mu},$$

and constraints of the preceding section. This model guarantees a balance of the operating budget at the end of the horizon, including deficits which have accumulated and including the sequence of subsidies over all districts and all time periods. Thus, at the end of the horizon the computed tax rate $\mu \geq \bar{\mu}$ yields:

$$(17) \quad \sum_{\theta} \sum_i \mu M_{\theta}^{(i)} = \sum_{\theta} \sum_i [C_{\theta}^{(i)} - F_{\theta}^{(i)}].$$

These models are to be contrasted against possibilities of district tax rates, μ_{θ} , such as the following.

District Tax Rates and Choice of Objectives

Weighted Sums of District Tax Rates. Consider minimizing the weighted overall rates. Let w_{θ} = weight for district θ . Then

the model is given by

$$(18) \quad \begin{aligned} & \text{Min } \sum_{\theta} \mu_{\theta} w_{\theta} \\ \text{s.t. } & \mu_{\theta} \geq (C_{\theta}^{(t)} - F_{\theta}^{(t)}) / M_{\theta}^{(t)}, \\ & C_{\theta}^{(t)} - F_{\theta}^{(t)} \geq 0 \\ & \mu_{\theta} \geq \bar{\mu} \end{aligned}$$

for all θ and all t together with the constraints of the preceding section. Various weight schemes can be devised including a system of weights which depends on variables in the model itself, e.g., tax revenue per student, market value per student, etc. Note that when $w_{\theta}=1$ for all θ , then the criterion is the minimization of the sum of the tax rates.

Minimizing Lexicographically the Maximum Tax Rate. Analogous to the developments of Charnes-Kortanek [4] on convex nuclei in n -person games we may introduce new game theoretic solution notions for the tax rates $\{\mu_{\theta} : \text{all districts } \theta\}$. For example, related to Schmeidler's nucleolus (see [4]), we may introduce the following ordering on the tax rates. For each selection of independent variables $\{N_{\theta i}^{(t)} : \text{all } \theta, i, \text{ and } t\}$ (the construction sequences) and funds subsidies $\{F_{\theta}^{(t)} : \text{all } \theta \text{ and } t\}$ a feasible set of μ_{θ} exists for the constraint set (18). For every such feasible set $\{\mu_{\theta}\}$ order these in decreasing order, i.e., find a permutation

$$\sigma: [1, 2, \dots, \bar{\theta}] \rightarrow [1, 2, \dots, \bar{\theta}]$$

$$\text{such that } \mu_{\theta\sigma(1)} \geq \mu_{\theta\sigma(2)} \geq \dots \geq \mu_{\theta\sigma(\bar{\theta})}.$$

Now ask is there a set of variables $\{N_{\theta i}^{(t)}\}$ and $\{F_{\theta}^{(t)}\}$ which yields a lexicographically smallest such ordering? This is analogous in the job scheduling context (see [12]) to minimizing maximum lateness, and the game concepts have had application there.

If we denote the set of all permutations

above by P_n , and if c is a large positive number, possibly non-Archimedian, the optimization problem for this is given by (see [4]):

$$(19) \quad \begin{aligned} & \text{Min } \nu \\ \text{s.t. } & - \sum_{\theta=1}^{\bar{\theta}} \mu_{\theta} c^{\sigma(\theta)} + \nu \geq 0 \\ & \text{for all } \sigma \in P_n \\ & \mu_{\theta} M_{\theta}^{(t)} + F_{\theta}^{(t)} \geq C_{\theta}^{(t)} \\ & - F_{\theta}^{(t)} \geq -C_{\theta}^{(t)} \\ & \text{for all } \theta \text{ and } t \\ & \mu_{\theta} \geq \bar{\mu}, \end{aligned}$$

and the constraints of the preceding section.

Finally one may analogously consider as in [4] any convex function of the $\{\mu_{\theta}\}$ as an objective, and if in particular the function is strictly convex then the set of μ_{θ} 's will be unique.

Concluding Remarks

Clearly, our models at this point are exploratory, but they appear to be among the first which couple up both individual units across all time periods, and thereby provide a generalization of the formulations of Charnes-Clower-Kortanek [3] on static coherent decentralization theory. In the case where nonlinearities are present, extensions of these results by Charnes-Fiacco-Littlechild [5] afford the kind of coherency desired and expressed in earlier sections of this paper.

The limitation of space prevents adequate comparison of our alternative models. However, it is obvious that different models may yield different results so that additional criteria of choice are needed. All that is claimed here is that these models represent a first step toward the analysis and understanding of the kinds of nontraditional problems associated with this non-market type of decentralization.

REFERENCES

1. Barr, James L., and Otto A. Davis, "An Elementary Political and Economic Theory of the Expenditures of Local Governments," *Southern Economic Journal* 33 (1966), 149-165.
2. Baumol, W. J., and T. Fabian, "Decomposition, Pricing for Decentralization and External Economies," *Management Science*, II No. 1, (September 1964).
3. Charnes, A., R. W. Clower, and K. O. Kortanek, "Effective Control Through Coherent Decentralization with Preemptive Goals," *Econometrica*, Vol. 35, No. 2 (April 1967).
4. Charnes, A., and K. O. Kortanek, "On Classes of Convex and Preemptive Nuclei for n -Person Games," Proceedings of the 1967 Princeton Symposium on Mathematical Programming, edited by H. W. Kuhn, 377-390, 1970.
5. Charnes, A., A. V. Fiacco and S. C. Littlechild, "Convex Approximants and Decentralization: A SUMT Approach," *Systems Research Memorandum No. 165*, Northwestern University, Evanston, Ill. (December 1966).
6. Charnes, A. and K. O. Kortanek, "On the Status of Separability and Non-Separability in Decentralization Theory," *Management Science: Applications*, Vol. 15, No. 2, (October 1968).
7. Davis, O. A., "Empirical Evidence of Political Influences Upon the Expenditure Policies of Public Schools," *The Public Economy of Urban Communities*, 92-111, Resources for the Future, Inc., Washington, D.C. (1965).
8. Davis, O. A., "Quality and Inequality: Some Economic Issues Related to the Choice of Educational Policy," *The University of Chicago Center for Policy Study*, The Univ. of Chicago (1969), 89-110.
9. Eastman, C. M. and K. O. Kortanek, "Modeling School Facility Requirements in New Communities," *Management Science*, 16 (1970), B784-799.
10. Hass, Jerome E., "Transfer Pricing in a Decentralized Firm," *Management Science: Applications*, Vol. 14, No. 6, (February 1968).
11. Hurwicz, Leonid, "On the Concept and Possibility of Informational Decentralization," *American Economic Review*, Vol. 59, No. 2, (May 1969), 513-524.
12. Kortanek, K. O. and W. L. Maxwell, "On A New Class of Combinatoric Optimizers for Multi-Product Single Machine Scheduling," *Management Science*, 15 (1969), 239-248.
13. Kortanek, K. O., "Effective Control Through Decentralization Separably and Non-Separably Structured Organizations," to appear in *Interorganizational Decision Making*, R. K. Chisholm, M. Radnor and M. F. Tuite (eds.), Aldine Publishing Company, forthcoming (1971).
14. Marschak, T., "Centralization and Decentralization in Economic Organizations," *Econometrica*, Vol. 27 No. 3, (July 1959).
15. Whinston, A., "Price Coordination in Decentralized Systems," Ph.D. thesis and O.N.R. Research Memo No. 99 (Carnegie Institute of Technology), (June 1962).
16. Whinston, A., "Price Guides in Decentralization Organizations," in W. W. Cooper, H. J. Leavitt, M. W. Shelley II (eds.), *New Perspectives in Organization Research*, J. Wiley & Sons, N.Y. (1964).
17. Report from The Mayor's Advisory Panel on Decentralization of the New York City Schools (1967).

DISCUSSION

CHARLES J. GOETZ: The Bradford-Oates paper rederived, in a public choice budgetary model, the theorem that matching grants are more stimulative than lump-sum grants of equivalent amount. The essentials of their analysis are correct, albeit perhaps more drawn-out than necessary. Somewhat paradoxically, the gist of my critique is that they claim both too little and too much for their analysis.

There are advantages in reframing their proof in simple opportunity-set terms. In switching from the matching to the lump-sum form, each voter in effect exchanges the triangular set of opportunities *FJC* in their Figure 5 for the set *FAH*. But the pivotal median voter and all inframedian voters are known to have rejected, under the matching grant, all other opportunities in favor of some point in *FAH*, namely one on line segment *AF*.

Thus, we can deduce that under the block grant, when the only loss is the "irrelevant" set *FJC*, all voters in a controlling coalition must again choose some point in *FAH*. In the highly improbable limiting case, the results of the matching and block grants would be identical; in all other cases, the matching grant would be more stimulative because the median voter finds some point in *FAH* preferable to the matching grant equilibrium *F*. This alternative proof does not require convexity of preferences, an assumption Bradford and Oates feel limits their results.

However, even the original proof offered by Bradford and Oates is not, as they claim, limited either to a single collective decision rule or to a single public good. If, for instance, a qualified majority rule requiring a two-thirds vote to alter the status quo were adopted, the alternative budget lines would merely be pivoted through the matching-grant optimum of the voter whose preference marks the minimum required coalition. The rest of the analysis follows as above.

With respect to the multipublic good model, a slightly modified version of the Hicks composite good theorem can be applied since we can isolate a subset of goods whose relative prices remain constant. The ordinate of Brad-

ford and Oates' Figure 5 can be interpreted as a composite consisting of the suboptimized equilibrium bundles of all goods, public and private, except the public good eligible for the matching-grant subsidy. Hence, the results of the paper are not really restricted to the case of a single public good.

It is from a methodological standpoint that I would fault the authors for implying *excessive* generality. As they acknowledge in the opening section, multiperson public choice models are designed to supplant the inappropriate, single-preference models of government decision making. Also, they allude to the Goetz-McKnew paper which derives contradictory results. If the Bradford-Oates and Goetz-McKnew models are each internally correct, as I believe they are, the choice between their contradictory results may be based on the appropriateness of the models. The Bradford-Oates model, as generalizable to *n* public goods, is Wicksellian in the sense that tax level and the mix of public goods are the subject of a single voting decision. The Goetz-McKnew model, by contrast, was explicitly designed to reflect so-called "general funds" budgeting in which an overall public budget is determined in one decision, while the relative allocations to individual public goods are determined in another decision. While I will not be dogmatic about which model more nearly fits most of the cases in the real world, you can guess the direction of my biases.

Finally, while this is in no sense a criticism, I should point out that the Bradford-Oates results are atypical of public choice models in the sense that they *are* able to validate a theorem derived in the single-preference models.

THOMAS E. BORCHERDING: Although I find Professor Hein's outline of the issues of fiscal decentralism perceptive and provoking, I must criticize his model on two methodological grounds. For one, his characterization of fiscal centralism seems incorrect. What he calls centralism would be better termed autarky. He assumes that initially under centralism local government goods and services are locally financed and decentralization then takes the

form of grants from the central government to these local units. This in turn requires an increase in the budget of the higher unit accomplished by a tax increase.

An alternative view would be to assume an initial position where some of the local government goods were financed and distributed by the central political unit. (Centralism exists either as a way of internalizing the inter-unit spillovers from the local public goods produced or as a method for redistributing income between units.) Decentralism would then be accomplished by a reduction of the latter unit's exhaustive expenditures on local government goods and a simultaneous transfer of these revenues to the noncentral governments. Since the relevance of comparative static predictions depends on the correct choice of assumptions about both the initial and final positions, the difference between Heins' and this alternative conception could be important.

My other concern is the model's total neglect of the mechanism linking the decisions of noncentral units with the central unit's decision as to the method of finance and dispersal of the grants-in-aid. Since Heins assumes that local government allocations are a function of citizens' preferences, consistency requires that he recognize that these same voters are also members of the central polity as well. Given this recognition, taking grants-in-aid as an exogenous variable seems wholly inappropriate. Furthermore, other interdependencies are also neglected. For instance, if decentralization involves the reduction of certain external decision-making costs (see Gordon Tullock, *Private Wants, Public Needs*, New York, Basic Books, 1970, Ch. 1) the marginal evaluation functions of local decisionmakers will shift even if the grant leads to no transfers of measured income.

One thing about the paper that I thought was particularly useful was his employment of a three-person, majority-rule model. Most of the other models that I have seen in this area treat the community as if it were an individual. This "orthodox" approach dismisses the reason for collective choice—the resolution of heterogeneous and conflicting preferences where unanimous choice is excessively costly. Heins' recognition of this is commendable.

However, he might have brought in the possibility of logrolling more directly, but this is of minor consequence. Had he done so he might have seen that the increase in "social tension" (Fig. 3-b) resulting from the introduction of grants-in-aid might well have been reduced or acerbated, depending on whether logrolling was, or was not efficient. If all groups affected by the chosen level of local government goods and their means of finance are party to the decision, excess demands after logrolling will be zero for each group. (This implies that logrolling approximates a perfect set of sidepayments. Thus, the marginal bribe necessary to change the allocation exceeds the value realized by this alteration.) Of course, incomes might be altered which might be a source of conflict. On the other hand, if some groups are not party to the decision (because of informational impedimenta or lack of political assets) logrolling might push them even further out of equilibrium.

Finally, given Heins' autarky-decentralization assumption, I think that more can be said about the aggregate "stimulative" effects of intergovernmental grants than is suggested. Neglecting the distributional effects of grants *within* local units and concentrating only on the effects *between* units, it can be demonstrated that the net effect of a movement from autarky to decentralization will be the expansion of total expenditures on local government goods and services. This hypothesis is based on the three following empirical propositions.

First, grants have only income effects. I think that most observers would concur with this assertion since grants are almost never open-ended, even though they are usually on a conditional and matching basis. In point of fact, most units spend in excess of that which they are eligible for in matching funds. Thus, the effects of most conditional grants are no different from unconditional aid, i.e., they are purely inframarginal. Second, virtually all empirical studies of noncentral governmental expenditures indicate positive, but less than unitary, income elasticities. Third, given Heins' assumption that these grants will be financed by a rise in the progressive income tax rate structure, the distribution of income will tend to shift in favor of the poorer units. (This assumes that the grants will not equal the addi-

tional taxes except for the middle-income units. The *net* grant residuum is negative for the wealthier units and is just equal to the positive grant residuum realized by the poorer units.)

The results of all this is an absolute increase in government expenditures of poorer local political units by an amount greater than the decrease in the expenditures of the wealthier units.

Note, however, that if the alternative conception of the centralism-decentralism problem is employed, the central unit's budget remains constant while its output of local public goods is decreased. The income change of a local unit then is the difference between the grants received and the value of the local government goods no longer supplied by the central unit. The net "stimulative" effects of grants in this context are not clear.

WALTER E. WILLIAMS: When we discuss the various arguments of fiscal decentralization we must first address ourselves to the question referring to the general goal(s) of the proposed decentralization. One goal of fiscal decentralization may seek to enhance economic efficiency while another may place emphasis on some sort of "ethically acceptable" redistributive goals, e.g., equal consumption of good X.

To the extent that these questions are kept in mind both the Heins and the Bradford & Oates papers raise some important questions regarding the impact of increased decentralization by the federal government in the form of matching and flat grants to sovereign governmental units. One of the authors' conclusions is that matching grants are more "stimulative" than flat grants. This means that matching grants tend to *stimulate* public expenditures more so than flat grants—a greater proportion of the latter may wind up as being tax rebates to the taxpayer.

So far as the models go, I have no objections. But as instruments of policy recommendation the models have a basic weakness. To help us discover this basic weakness let us go a step further to disaggregate the grant category. One such disaggregation is *specific* flat grants and *general* flat grants on

the one hand and on the other *specific* matching grants and *general* matching grants. A specific grant (matching or nonmatching) is just as the name implies—the granting agency specifies the use of the funds, and desired redistributive objectives can be somewhat assured, e.g., welfare programs, remedial education, etc. A nonspecific grant has no such assurances, and its use depends upon the preferences of the receiving community.

Having noted these points, while not necessarily taking the strict assumptionist position, I question the *usefulness* of the Heins and Bradford-Oates models where the recipient merely receives the flat grant or matching grant. In fact if we relax the model to allow for a more descriptive assumption of choice (on behalf of the grant recipients), we may derive a "stimulative" effect in the aggregate quite different from that predicted in the Heins paper and the Bradford-Oates paper.

Suppose, for example, we have a mayor of a large metropolitan area. Suppose further that the central fiscal authority offers periodically flat grants (specific and general) and matching grants (specific and general). Let the specific grants be for projects determined by the central fiscal authority. The utility function of the mayor has the following arguments:

$$(1) \quad U_m = U_m(Y_m, SC, Prob-re, PEP, x_1, x_2, \dots, x_n)$$

where:

U_m = utility of the mayor

Y_m = income of the mayor

SC = the size of the city

$Prob-re$ = the probability of re-election

PEP = public expenditure projects 1, 2, 3, . . . p

x = private goods 1, 2, 3, . . . n

The mayor maximizes utility subject to the following simplified constraint:

$$(2) \quad \begin{aligned} &\text{Probability of re-election } (Prob-re) \\ &= f(TR, PEP) \end{aligned}$$

where TR is the tax rate.

Let the first case be the receipt of a flat

grant (general). Such a grant will leave the mayor (and/or his officers) with the freedom of choice as to the disposition of the funds so far as public expenditure projects to choose. Under these circumstances the official will choose the public expenditure project which maximizes his utility. As we have seen, one argument in his utility function is the probability of his being re-elected; a higher probability of re-election may be associated with his choice to subsidize public expenditure project number three (PEP_3) than (PEP_4). That is:

$$(3) \quad \frac{\partial U_m}{\partial \text{Prob-re}} > 0 \quad \text{and,} \\ \frac{\partial \text{Prob-re}}{\partial (PEP_3)} > \frac{\partial \text{Prob-re}}{\partial (PEP_4)}$$

and therefore:

$$(4) \quad \frac{\partial U_m}{\partial (PEP_3)} > \frac{\partial U_m}{\partial (PEP_4)}$$

This being the case we can expect the mayor to select (PEP_3), say higher education, as the project on which to spend the grant funds.

Let the second case be the receipt of a matching grant. Say that the matching requirement is 1:1. The mayor upon receipt of the offer will attempt to maximize utility. But note that in the case of the matching grant he has to make a decision with respect to raising taxes (or reshuffle existing tax expenditures—but let us assume that he cannot exercise this option). Raising taxes in order to meet the matching requirement is one of the postulated constraints on his probability of re-election, i.e., $\partial \text{Prob-re} / \partial TR < 0$. Thus given that a rise in taxes may reduce the probability of re-election the mayor *may* decline the assistance of the matching grant. Exceptions to our mayor's choice to refuse the matching grant may be the cases where the matching grant (unspecified) allows him the option to choose (PEP_k) which has significant benefits to the dominant coalition of voters—(PEP_k) could be parks in middle and high income suburbs. Therefore, if:

$$(5) \quad \left| \frac{\partial \text{Prob-re}}{\partial TR} \right| < \left| \frac{\partial \text{Prob-re}}{\partial (PEP_i)} \right|$$

the mayor will *accept* the grant when the i^{th} project is one expenditure option available to him. But if the matching grant (specified) is for school lunches in a poverty community (PEP_p), the mayor may decline the matching-grant offer from the central fiscal authority because the dominant coalition, while experiencing an increase in taxes, may not perceive significant benefits from (PEP_p) and therefore:

$$(6) \quad \left| \frac{\partial \text{Prob-re}}{\partial TR} \right| > \left| \frac{\partial \text{Prob-re}}{\partial (PEP_p)} \right|.$$

In sum, what is being said here is not offered as refutation of the hypothesis advanced by the Heins paper and the Bradford-Oates paper. Instead I question the practical usefulness of the model for policy recommendation. The problem, I think, is the implication that the central fiscal authority has a task of deciding whether to give a matching grant or a flat grant and he has to decide which grant will have a greater "stimulative" effect. But as I have argued, if the receiving governmental unit has a choice to accept or reject grants, then depending on the *mix* of specified matching grants (kinds of projects on which the grant may be spent) and community preferences, we may find that in the *aggregate* flat grants may stimulate more public expenditures than matching grants.

Turning to the Davis-Kortanek paper, much of the recent controversy regarding the decentralization of the public schools has to do with the dissatisfaction by parents and other community elements with the kind of education being received by students. The argument basically boils down to the desire of parents for the public school system to have greater flexibility such that local schools reflect the preferences of the local community. To the extent that this controversy exists, alternative institutional arrangements in the delivery of public education is an important area for political economists to study.

The Davis-Kortanek paper is one such at-

tempt to give scholarly attention to this problem. The authors point out the stated objectives of New York City's Mayor's Advisory Panel. They point out, too, the obvious difficulties of choosing those (recommendations) as objective functions. Also the authors reject the approach of orthodox welfare economics which claims that the objective is to maximize an "ethically acceptable social welfare function."

Instead the authors choose mathematical programming as their analytic tool and say that they will specify several alternative objective functions—each of which will serve as a surrogate for a social welfare function. One such objective function given is to provide a *unit need*, whose objective measurement by the authors is the cost of accommodating one student of age i during time period t in school district θ . $C_{\theta}^{(i)}$ is the total cost of instruction in district θ . Later on in their paper Davis and Kortanek seek to minimize a single tax rate (μ) across all districts versus differing tax rates (μ_{θ}) for individual districts, subject to conditions that the tax is equal to or greater than the positive difference between cost of instruction and state subsidy divided by the market value of the houses in the district(s).

This is all very interesting and as far as their arguments go I can hardly find much

exception to them—but I cannot see how these arguments really come to grips with the issue of decentralization as suggested by the title of the Davis-Kortanek paper. Further, they give us only one objective function, which implies that we minimize the cost of one good (education) and maximize the other ($n-1$) goods.

When one thinks of the elements of the debate between centralization and decentralization, one mostly thinks of how alternative institutional arrangements will allow us to get more from given resources, get more complete information, penalize decision makers for "wrong" choices, and reward them for "correct" choices. The authors did not tell us anything along these lines. Moreover they never came to grips with the dimension of educational output which seems to be a crucial element in the whole controversy.

In conclusion, I might point out that the development of models including the sequencing of housing, the district student-growth function, and the interdependencies between school districts is encouraging. This is very important because it forces us to recognize that it is by no means certain that decentralization will produce the optimal outcome when there is full acknowledgement of the interrelationships between sectors and between periods.

AMERICAN ECONOMIC ASSOCIATION

PROCEEDINGS
OF THE
EIGHTY-THIRD
ANNUAL
MEETING

DETROIT, MICHIGAN
DECEMBER 27-30, 1970

PROCEEDINGS OF THE AMERICAN ECONOMIC ASSOCIATION

ANNUAL BUSINESS MEETING DECEMBER 29, 1970 COBO HALL, DETROIT, MICHIGAN

The Eighty-third Annual Business Meeting of the American Economic Association was called to order by President Wassily Leontief in Room 2040 of Cobo Hall, Detroit, Michigan, at 9:30 p.m. on December 29, 1970.

It was VOTED to approve the minutes of the Annual Business Meeting of December 29, 1969, and to ratify the minutes of the Executive Committee meetings held on March 7-8, 1969, and December 27, 1969, as published in the *American Economic Review*, May 1970, pages 487-93.

The Secretary presented the report of the Committee on Elections and the certification of the new officers for 1970 as follows:

In accordance with the bylaws on election procedure, I hereby certify the results of the recent balloting and present the reports of the Nominating Committee, the Electoral College, and the Committee on Elections.

The Nominating Committee, consisting of Milton Friedman, Chairman, Irma G. Adelman, Andrew F. Brimmer, Hollis B. Chenery, Albert Fishlow, and Harold M. Somers, submitted the nominations listed below for Vice-Presidents and Executive Committee. The Electoral College, consisting of the members of the Nominating and Executive Committees meeting together, selected the nominee for President-elect.

President-elect

John Kenneth Galbraith

Vice-Presidents

James M. Buchanan
Franco Modigliani
Henri Theil
Clifton R. Wharton, Jr.

Executive Committee

Armen A. Alchian
Robert Eisner
Carl F. Christ
John R. Meyer

The former Secretary, Harold F. Williamson, prepared biographical sketches of the candidates and distributed ballots in September. The Committee on Elections, consisting of James S. Worley, Chairman, William O. Thweatt, and Rendigs Fels, canvassed the ballots and filed the results with the Secretary. From the Report of the Committee on Elections, I have the following information:

Number of envelopes without name for identification	331
Number of envelopes received too late	68
Number of defective ballots	0
Number of legal ballots	6,278
	<hr/> 6,677

On the basis of the canvass to the votes cast, I certify that the following persons have been duly elected to the respective offices:

President-elect (for a term of one year)

John Kenneth Galbraith

Vice-Presidents (for a term of one year)

James M. Buchanan
Franco Modigliani

Members of the Executive Committee (for a term of three years)

Robert Eisner
John R. Meyer

In accordance with the action of the Executive Committee at its meeting of March 6-7, 1970, an amendment to Section IV, Paragraph 2, of the bylaws was submitted to the members in a mail ballot in conjunction with the balloting for officers canvassed by the Committee on Elections. On the basis of the canvass, I certify that the amendment was approved. As amended, Section IV, Paragraph 2, now reads:

Before October 1 of each year, the President-elect of the Association shall appoint a Nominating Committee for the following year, this Committee to consist of a past officer as Chairman and not less than five other members of the Association. In addition to appointees chosen by the President-elect, the Committee shall include any other member of the Association nominated by petition including signatures and addresses of not less than 2 percent of the members of the Association, delivered to the Secretary before December 1. No member of the Association may validly petition for more than one nominee for the Committee.

The names of the Committee shall be announced to the membership immediately following its appointment and the membership invited to suggest nominees for the various offices to the Committee. The Nominating Committee for each year shall be instructed to present to the Secretary of the Association on or before March 1 a nominee for the President-elect and two or more nominations for each other elective office to be filled, except the presidency, all these nominees being members of the Association. The members of the Nominating and Executive Committees shall constitute an Electoral College which shall consider the nominee of the Nominating Committee for the President-elect and select a single candidate for that office. In the voting in the Electoral College the Nominating Committee, whatever its number of members, shall not have greater weight than the Executive Committee.

The Secretary shall inform all members of the Association of the actions of the Nominating Committee and the Electoral College not later than the June issue of the *American Economic Review*. An additional nomination for any office may be made by petition, delivered to the Secretary by August 1, including signatures and addresses of not less than 6 percent of the membership of the Association for the office of President-elect and not less than 4 percent for each of the other offices. No member of the Association may validly petition for more than one nominee for the Executive Committee, one nominee for Vice-President, and one nominee for President-elect. The election of officers by the membership shall take place by a mail ballot conducted by the Secretary before December 1 each year.

The ballot shall list all nominees alphabetically, with indication "nominated by petition" where applicable. Space shall be provided on the ballot for the individual voter's alternative choices for all offices. The candidates with the highest number of votes for the various offices will be elected.

The President-elect may, at his discretion and with the advice and consent of the Executive Committee, appoint a Program Committee for the annual meeting of the succeeding year.

Reports submitted by the Secretary, the Treasurer, the Managing Editor of the *American Economic Review* and the Managing Editor of the *Journal of Economic Literature* were circulated in written form and accepted. (These reports are published below.)

At this point, President Leontief called on Harold F. Williamson to recognize Miss Gertrude Tait and express to her the appreciation of the Association for her forty-five years of service. Mr. Williamson announced that the Association was giving her an appropriate gift. As Miss Tait left the platform, the members rose to their feet and applauded.

President Leontief then introduced the new President, James Tobin, who took over the chair. President Tobin read the names of the members of the Nominating Committee for 1971, which had been published in the *American Economic Review* for December 1970.

President Tobin introduced President-elect John Kenneth Galbraith.

The Treasurer moved adoption of an amendment to Section I, Paragraph 2, of the bylaws that had been approved by the Executive Committee at its meeting of March 6-7, 1970. The effect of the proposed amendment was to double the dues for all classes of members and to delete the word "consecutive" from the provision governing junior members. The motion was seconded. It was voted to divide the motion and take separate votes on each

class of membership. The subsequent voting adopted all the provisions of the amendment, including deletion of "consecutive," except the proposed change in the dues of junior members. The new wording of Section I, Paragraph 2, as adopted, reads as follows:

There shall be six classes of members other than honorary: members paying an annual fee of \$20.00; family members (two or more living at the same address, second membership without subscription to the publications of the Association) paying an annual fee of \$2.00; junior members (available to registered students for three years only) paying an annual fee of \$5.00; subscribing members paying an annual fee of \$24.00; contributing members paying an annual fee of \$50.00 or more; and life members comprising those members who contribute \$400.00 or more in a single payment. Life members shall be exempt from annual fees. Members shall have each year the privilege of designating the class of membership they choose for that year.

It was then VOTED that the Executive Committee be empowered to raise the dues of junior members to the minimum (if any) necessary to retain second class mailing privileges.¹

Mr. Elliott R. Morss moved that the Association endorse the idea of establishing a secretariat in Washington, D.C., to increase the communications between Federal policymakers and Association economists but *not* to take policy positions. The motion was seconded. It was VOTED first to substitute for Mr. Morss's motion and then to adopt a motion that the President appoint a committee to consider the proposed secretariat and to report its findings to the Executive Committee and to the next annual business meeting.

Subsequently, it was VOTED to adopt Mr. Morss's original motion.

Mr. Morss moved that the President appoint a committee to investigate the desirability of the Association establishing a new journal devoted to economic policy. After the motion was seconded, it was put to a vote and lost.

Mr. Galbraith then asked that the minutes record his intention in testifying before Congressional committees and carrying on other political activity to make it clear that he is not speaking for the Association but is acting as an individual.

A resolution was offered proposing that there be a session at the 1971 convention on the role of women in economic life. The resolution was withdrawn as superfluous after Mr. Galbraith assured the members that such a session was already being planned. Mr. Galbraith solicited specific suggestions for the subjects of papers and for possible authors.

The Secretary presented the following resolution, which was adopted:

This meeting would like to record a vote of special thanks to Dean E. Richardson, Chairman; Richard H. Hanson, Executive Secretary; Patricia Shontz and S. William Takai, Vice Chairmen; and the other members of the 1970 National Convention Committee for their efficient management of the annual meeting of the Allied Social Science Associations.

The Secretary presented the following resolution, which was adopted by acclamation:

Be it resolved that this meeting commend James Tobin and members of the Program Committee for planning an outstanding program for the 1970 Annual Meeting.

The meeting was adjourned at 11:45 p.m.

¹ In accordance with this vote, the dues of junior members were subsequently raised to \$10.00.

MINUTES OF THE EXECUTIVE COMMITTEE MEETINGS

1. Minutes of the spring meeting held in New York City, March 6-7, 1970:

The first meeting of the 1970 Executive Committee was called to order at 9:30 a.m. in the Americana Hotel, New York, New York, March 6, 1970. The following were present: Wassily Leontief, presiding, G. H. Borts, Mary Jean Bowman, E. D. Domar, Robert Dorfman, Otto Eckstein, W. J. Fellner, A. C. Harberger, Mark Perlman, C. L. Schultze, Tibor Scitovsky, James Tobin, and H. F. Williamson. Absent were: K. E. Boulding and Anthony Scott. Present as members of the Nominating Committee were: Milton Friedman, Chairman, Irma G. Adelman, A. F. Brimmer, H. B. Chenery, Albert Fishlow, and H. M. Somers. Present as guests were: Marcus Alexis, Robert Bell, Robert Brown, Rendigs Fels, Harriet Harper, J. P. Miller, T. H. Spratlen, Henry S. Terrell, and C. Z. Wilson.

1. *President's Remarks* (Wassily Leontief). President Leontief outlined the items on the agenda to be covered at the meetings.

2. *Minutes*. It was VOTED to approve the minutes of the December 27, 1969, Executive Committee meeting as corrected.

3. *Report of the Secretary* (H. F. Williamson). The Secretary reported that total registration for the annual meeting in New York City was approximately 7,000. It was indicated that the chairman of the Committee on Local Arrangements for the 1970 Detroit meeting had not yet been appointed, but it was hoped that the Committee would be organized in the near future. Following a discussion of future meeting plans, it was VOTED to hold the 1975 annual meeting in Dallas, Texas, and the 1976 annual meeting in Atlantic City, New Jersey.

Papers and Proceedings. It was announced that 28,000 copies of the May 1970 *Papers and Proceedings* had been ordered and that the total manuscript pages was 1,546. The Secretary noted that, at the request of the Black Caucus, a report of the informal meeting of the Black Caucus and A.E.A. representatives and the "Statement of Concern of the Caucus of Black Economists" presented at the meeting held in New York City on the eve-

ning of December 28, 1969, was included in the 1970 *Papers and Proceedings*.

The National Register of Scientific and Technical Personnel. It was announced that 23,529 questionnaires had been mailed to members of the profession; the expectation was that at least 13,000 would be returned.

Handbook. It was reported that plans were under way to distribute the 1969 *Handbook* within the next few weeks.

4. *Auditor's Report* (H. F. Williamson). Copies of the 1969 Auditors' Report were distributed. The Treasurer called special attention to the fact that the Association's total income for the year ending December 31, 1969, was \$373,470, total expenses were \$482,282, and expenses exceeded income for the year by \$108,812, and that the unrestricted balance as of December 31, 1969, was \$305,937.

5. *Report of the Finance Committee* (C. W. Farnham). In the absence of the Chairman, H. F. Williamson reported informally that in view of the current unsettled state of the financial markets the Finance Committee was following a cautious investment policy.

6. *Report of the Treasurer* (H. F. Williamson). The Treasurer reviewed the financial status of the Association, including the prospects for 1970. A rough estimate of income and expenditures strongly suggested that the excess of expenses over income for 1970 would be close to \$150,000. It is anticipated that most of the increase would be the result of higher printing costs, salary adjustments, and the expenses involved in moving the office of the Secretary-Treasurer from Evanston. Because it seemed quite likely that the Association would have to borrow money to cover part of the expenses for 1970, it was VOTED to authorize any two of the three following to arrange for a bank loan: the President, the Secretary-Treasurer, and the Chairman of the Finance Committee.

Following the discussion of the future financial needs of the Association, it was VOTED to rescind the action taken at the December 1969 Executive Committee meeting regarding annual dues and to recommend to the December 1970 Business Meeting that regular annual dues and subscriptions be

raised to \$20, with appropriate adjustments in the other classes of members, to be effective January 1, 1971.

7. *Report of the Editor, A.E.R.* (G. H. Borts). The Editor of the *Review* outlined his 1970 budget and presented an estimated budget for 1971. Following a discussion of the problem of reducing the time lag between the receipt of articles and their acceptance or rejection for publication, it was VOTED to authorize the Editor to spend up to \$2,500 in addition to the amount in his authorized budget for 1970 for editorial services.

8. *Report of the Editor, J.E.L.* (Mark Perlman). The Editor reported informally on his plans for future issues of the *Journal* and presented a budget estimate for 1971. He also noted that Mrs. Koopmans expected to start work in the near future on the 1968 volume of the *Index of Economic Journals* and that the *J.E.L.* was working on a two-year volume of the *Index of Economic Journals* for 1969-70.

9. *Reports of Committees.*

Committee on Research and Publications (J. P. Miller). In his report, the Chairman noted that after a full discussion of the status of the Pareto translation and plans for the future, the Committee had agreed that the following information and recommendations should be presented to the Executive Committee:

1. The Association should go ahead with the publication of the series of readings as originally planned by Professor Chipman. Over and beyond the expenses already incurred in the payment of royalties to Droz and for translation to Professor Dehem, it is anticipated that there will be about \$4,000 additional expenses to be incurred for revision of the translation by Mr. Cairncross and incidental expenses, including xeroxing, postage, telephone, etc., by Mr. Chipman. It was agreed that he will include in the volume the 133-page Appendix to the French edition of the *Manuel*.

2. Professor Chipman has been negotiating with Mr. Allen, of the Pareto Fund, for the publication of another volume of selections from Pareto. It is understood that Mr. Allen, speaking for the Pareto Fund, is prepared to support this book of essays. We propose that the American Economic Association cosponsor this and publish it in the same series with Volume I with the understanding that the Pareto Fund and Mr. Allen will support at least all costs other than

publication costs. We would ask the Executive Committee to agree to support publication costs through the usual arrangements with Irwin.

There is a question as to who should hold the copyright on this. With respect to Volume I, the American Economic Association seems to have acknowledged the copyright position of Droz, but Mr. Allen is unwilling to recognize their right or support any payment to them. He alleges they have no copyright position with reference to Pareto's publications. We propose, therefore, to pursue this matter of copyright with the Pareto Fund and the Association lawyers. If it can be established that Droz does not have a copyright, we propose that the copyright be held, as usual, by the A.E.A. If this is not established satisfactorily, we propose that the copyright to this volume be held by Mr. Chipman whom the Pareto Fund has indicated they would support in case any legal actions are brought by Droz.

3. Although the Executive Committee at its December meetings suggested that we pursue the advisability of publishing the *Manuel*, we suggest that no action be taken on this at this time. We propose that the matter be reviewed after the copyright position with reference to Pareto's works has been clarified (as indicated under 4 below) and after we have had an opportunity to establish either that Kelley is not going to publish his edition of the *Manuel* or that the quality of the volume published by Kelley is unsatisfactory.

4. We think it is important to clarify the copyright problems. This would be essential if we should at some later date undertake to publish the *Manuel*. But it is also relevant to our arrangements with the Pareto Fund on the publication of the second volume of *Essays* edited by John Chipman. Consequently we propose the following steps: (a) Mr. Chipman pursue with Mr. Finer at Manchester, England, his findings concerning the copyright. Finer has published translations of several Pareto items. (b) That the Association's lawyer in Washington, D.C., Mr. Rubin, be asked to cooperate with lawyers for the Pareto Fund in exploring the Pareto copyright problem.

Following a discussion, it was VOTED to accept the foregoing recommendations of the Committee on Research and Publications.

It was further VOTED to accept the sug-

gestion of the Chairman that a new Committee on Research and Publications be appointed by the President to consider the future research and publication objectives of the Association. It was also VOTED to express the Association's gratitude to the members of the present Committee on Research and Publications for their services.

Committee on the Education and Training of Minority Group Economists (A. F. Brimmer). The Chairman reviewed the work of the Committee since its appointment following the annual meeting in New York. He outlined what appeared to be major problems in providing training opportunities for minority group economists, and he noted the following suggestions as to how the situation might be improved:

1. The Association should sponsor a new institution whose main assignment would be to help broaden opportunities for the training of Negro and other minority group economists.

2. The Executive Committee should arrange to have detailed plans for the new institution drawn up.

3. The new institution should serve—at least initially—as a secretariat. It would function as a clearinghouse for information on graduate training opportunities and as a means of identifying promising candidates. It might also sponsor a variety of programs itself.

4. The secretariat (or new institution) should be established as an adjunct to an existing educational institution—preferably a university with a strong graduate program in economics. Ideally, it should be located on the East Coast, but this decision should be governed by the availability of persons to direct the work and the willingness of the university to serve as host.

5. The possibility of establishing a separate teaching institution (modeled on the Economics Institute at the University of Colorado) was weighed by the Committee. However, because of the enormous financial and personnel implications of such a proposal, the Committee decided that the matter should be reserved for consideration by the new institution which is being recommended.

6. The financial requirements of the recommended program will vary greatly, depending on the approaches finally adopted

by the new institution. However, a rough estimate suggests an annual cost ranging from \$300,000 to \$500,000.

7. Moreover, the Committee believes the program should not be launched unless funding can be assured for at least three years. Thus, the required commitment would range from about \$1,000,000 to \$1,500,000.

8. These estimates are based on the assumption that the program, in its early years, should aim at increasing the enrollment of minority group economists seeking Ph.D. degrees by 50–75 per year.

9. Clearly, this long-term funding would have to be provided by foundations, the federal government, and by the participating universities. Developing this support would be one of the chief functions of the new institution.

10. In the meantime, it is recommended that the Association allocate about \$15,000 to enable the new institution to be formed.

11. Finally, the Committee on the Education and Training of Minority Group Economists might be asked to continue to serve as the Association's instrument for developing the program.

Following the presentation, it was VOTED to continue the Committee and to authorize an expenditure of not more than \$10,000 to cover this Committee's expenses. It was further VOTED that applications for obtaining outside financial support should be approved by the A.E.A. Executive Committee by mail ballot prior to their submission to foundations or government agencies.

Nominating Committee (Milton Friedman). The Executive Committee and the Nominating Committee met as an Electoral College to consider the nominees for the office of President-elect for 1971. After discussion, the nominee was selected and his acceptance obtained. Nominees for the other offices were discussed. The Chairman also submitted the recommendation of the Committee on Honors and Awards for the Distinguished Fellow award. The recipient was selected as the result of a ballot of the members of the Electoral College.

10. *Unfinished and Miscellaneous Business.*

On behalf of the Committee on the Selection of a New Secretary-Treasurer, W. J. Fellner reported that Rendigs Fels of Vanderbilt

University had been offered the appointment as Secretary-Treasurer of the Association and negotiations were being carried on with Vanderbilt University regarding the space for the Secretary's Office. It was agreed that W. J. Fellner should complete final arrangements with Professor Fels and Vanderbilt University for moving the office.

A committee representing the Black Caucus presented a prospectus, "Increasing the Supply of Black Economists." Following a discussion, it was agreed that the Association would endorse an application by the Black Caucus committee to the Ford Foundation for financial support to organize a series of conferences on the general subject of economic development of ethnic minorities, and it was VOTED to appoint the following A.E.A. committee to advise the Black Caucus on implementing the proposed program: James Tobin, Chairman, Charles Schultze, George Borts, Robert Dorfman, and Marcus Alexis. It was also VOTED to authorize an expenditure of up to \$10,000 to cover the expenses of this committee.

As a means of insuring greater representation of members in the selection of candidates for offices in the Association, James Tobin proposed that Section IV, Paragraph 2, of the Bylaws be changed to read as follows:

Before October 1 of each year, the President-elect of the Association shall appoint a Nominating Committee for the following year, this Committee to consist of a past officer as Chairman and not less than five other members of the Association. In addition to appointees chosen by the President-elect, the Committee shall include any other member of the Association nominated by petition including signatures and addresses of not less than 2 percent of the members of the Association, delivered to the Secretary before December 1. No member of the Association may validly petition for more than one nominee for the Committee.

The names of the Committee shall be announced to the membership immediately following its appointment, and the membership invited to suggest nominees for the various offices to the Committee. The Nominating Committee for each year shall be instructed to present to the Secretary of the Association on or before March 1 a nominee for the President-elect and two or more nominations for each other elective office to be filled, except the presidency, all these

nominees being members of the Association. The members of the Nominating and Executive Committees shall constitute an Electoral College which shall consider the nominee of the Nominating Committee for the President-elect and select a single candidate for that office. In the voting in the Electoral College the Nominating Committee, whatever its number of members, shall not have greater weight than the Executive Committee.

The Secretary shall inform all members of the Association of the actions of the Nominating Committee and the Electoral College not later than the June issue of the *American Economic Review*. An additional nomination for any office may be made by petition, delivered to the Secretary by August 1, including signatures and addresses of not less than 6 percent of the membership of the Association for the office of President-elect and not less than 4 percent for each of the other offices. No member of the Association may validly petition for more than one nominee for the Executive Committee, one nominee for Vice-President, and one nominee for President-elect. The election of officers by the membership shall take place by a mail ballot conducted by the Secretary before December 1 each year. The ballot shall list all nominees alphabetically, with indication "nominated by petition" where applicable. Space shall be provided on the ballot for the individual voter's alternative choices for all offices. The candidates with the highest number of votes for the various offices will be elected.

The President-elect may, at his discretion and with the advice and consent of the Executive Committee, appoint a Program Committee for the annual meeting of the succeeding year.

Following the discussion, it was VOTED to submit the proposed amendment to the members by mail ballot.

At the request of Anthony Scott, it was VOTED to conduct a mail ballot asking members to indicate their preference for continuing to hold our annual meeting between Christmas and New Year's or, alternatively, just prior to or immediately after Labor Day.

President-elect James Tobin discussed his plans for the program for the 1970 annual meeting. He recommended that a volume of abstracts of all main papers to be presented at

the meeting of the Allied Social Science Associations be prepared and distributed by mail prior to the meeting to everyone who preregisters for the meeting. He also recommended that facilities be made available at the meeting to duplicate papers for a fee at the request of members.

The matter of permission to reprint articles from the *American Economic Review* and the *Journal of Economic Literature* was discussed. It was VOTED that the recommended fee to the author for permission to reprint should be raised from \$100 to \$150.

It was VOTED to accept the motion by Robert Dorfman that the Executive Committee record its deep concern with regard to the breach of good order at the 1969 Business Meeting.

The meeting adjourned at 12:30 p.m. on March 7, 1970.

2. Minutes of the meeting held in Detroit, Michigan, December 27, 1970:

The second meeting of the 1970 Executive Committee was called to order at 2:15 p.m. in the Sheraton-Cadillac Hotel, Detroit, Michigan, December 27, 1970. The following members were present: Wassily Leontief, presiding, George H. Borts, Kenneth E. Boulding, Mary Jean Bowman, Evsey Domar, Robert Dorfman, Otto Eckstein, William J. Fellner, Rendigs Fels, Arnold Harberger, Mark Perlman, Charles L. Schultze, and James Tobin. Absent were: Tibor Scitovsky and Anthony Scott. Present as guests were: Marcus Alexis, Vittorio A. Bonomo, Andrew F. Brimmer, James M. Buchanan, Robert Eisner, C. Wells Farnham, John Kenneth Galbraith, Fritz Machlup, John R. Meyer, Charles T. Reich, Thaddeus H. Spratlen, Harold F. Williamson, and Charles Z. Wilson.

1. *Minutes.* The minutes of the meeting of March 6-7, 1970, were corrected and approved.

2. *Report of the Secretary* (Rendigs Fels). The Secretary reported that the referendum on the time of the annual meeting showed a preference of less than six to five for dates immediately before or after Labor Day rather than immediately after Christmas. He pointed out that firm commitments have been made for annual meetings after Christmas through 1976. In view of the difficulty of determining what the preferences of the membership really are, it was VOTED to experiment with holding the meetings at different times. An effort

will be made to shift the 1972 meeting scheduled for Toronto to a date in the fall and similarly for 1975 (Dallas) and 1976 (Atlantic City). The Committee was evenly divided between September and October. It was VOTED to give Harold F. Williamson the title of National Convention Coordinator and to give him a free hand in making such changes.

It was VOTED to approve the following statement of policy governing minutes submitted by the Secretary:

At the meeting of the Executive Committee of the American Economic Association on March 6, 1970, Mr. Tobin suggested that the future Secretary be asked to present to the Committee a statement on policy governing what is included in the minutes. Mr. Tobin's suggestion was prompted by discussion of the wisdom of including in the minutes of the Annual Business Meeting held on December 29, 1969, the statement of Mr. MacEwan attacking the Association and by Mr. Tullock objecting to the disturbance of the meeting by Mr. MacEwan and others.

The principal object of the minutes is to record actions taken. Inclusion of anything more, such as the statements of Messrs. MacEwan and Tullock or summaries of the remarks of other speakers, should be governed by the sole criterion of whether enough people will want to read it to make inclusion worth the cost.

Three implications of this criterion may be pointed out. First, the concept of equal time is irrelevant. If Mr. MacEwan's statement interests readers but Mr. Tullock's does not (or vice versa), then only the one should be printed. Second, editorial judgment must be exercised, normally by the Secretary, but occasionally by the Executive Committee when its judgment differs from his. Judgment calls are likely to be disputed by those adversely affected, as every umpire and referee knows, but that is no reason for shrinking from making them. Third, the minutes normally can be expected to include actions only. A speech or statement has to be exceptional to warrant the conclusion that its general interest makes it worth the cost.

A discussion of the *Papers and Proceedings* issue of the *American Economic Review* was generally favorable toward continuing publi-

cation. The question of whether to continue publication of abstracts of the papers before the annual meeting was postponed.

The Secretary reviewed other activities of the Association, which are treated more fully in the Secretary's Report printed below. It was VOTED to approve the Report.

3. *Report of the Treasurer* (Rendigs Fels). The report of the Treasurer, covering the calendar year 1969 (see below), was summarized. The Association had a deficit of \$108,812 in that year. The Treasurer pointed out that the deficit for 1970 would be considerably larger. In view of the pending proposal to double the dues and the uncertainty as to the price elasticity of demand for membership in the Association, the deficit or surplus for 1971 was hard to project. The Treasurer presented four different projections based on different assumptions with respect to the dues change and the elasticity of response. The most reasonable assumption about elasticity implies a further accounting deficit in 1971 but not a cash deficit. The difference between the two results from the investment income formula used by the Association.

The Treasurer reviewed the change in Section I, Paragraph 2, of the bylaws approved by the Executive Committee at its meeting of March 6-7, 1970, and recommended a further change, namely, deletion of the word "consecutive" in the provision concerning junior members. It was so VOTED. It was further VOTED that if the amendment to Section I, Paragraph 2, were rejected by the membership at the Annual Meeting on December 29, 1970, it would be submitted to the entire membership by mail ballot.

It was VOTED to authorize the Treasurer to open a deposit account at the First American National Bank of Nashville and/or the Third National Bank of Nashville and that any authorization granted to borrow money on behalf of the Association be extended to permit borrowing from either of said banks.

It was VOTED to approve the report of the Treasurer.

4. *Report of the Finance Committee* (C. Wells Farnham). The Chairman stated the belief of the Committee that the Association may properly borrow money to hold bonds, provided that rates of interest are favorable, but that it should not borrow money for the purpose of holding stocks. It was VOTED to continue, without limit of time, the borrowing authorization approved at the meeting of the

Executive Committee, March 6-7, 1970.

It was VOTED to elect Beryl Sprinkel to the Finance Committee. It was VOTED to empower the President to make an additional appointment to the Committee.

Mr. Farnham expressed the desire to resign from the Committee as soon as convenient. The Executive Committee expressed to him its deep appreciation for his long and outstanding service to the Association.

5. *Report of the Editor of the American Economic Review* (G. H. Borts). The Editor reviewed the operation of the Editorial Office of the *American Economic Review*, covered in full in the report printed below. It was VOTED to appoint Phillip D. Cagan, Daniel L. McFadden, Herbert Mohring, and Edmund S. Phelps to the Board of Editors and to approve the report of the Managing Editor.

6. *Report of the Editor of the Journal of Economic Literature* (Mark Perlman). The Editor reviewed the operations of the Editorial Office of the *Journal of Economic Literature*, covered in full in the report printed below. It was VOTED to appoint Irma G. Adelman, William J. Baumol, and Marshall Hall to the Board of Editors and to approve the report of the Managing Editor.

7. *Committee on the Education and Training of Minority Group Economists* (Andrew F. Brimmer). The Chairman submitted a written summary of recommendations. A full report is to be submitted in about a month. It was agreed that the full report would be reviewed by an Advisory Committee and considered by the Executive Committee at its March meeting. President Leontief announced that he was appointing as members of the Advisory Committee Robert Dorfman, Chairman, Marcus Alexis, G. H. Borts, Charles L. Schultze, Thaddeus H. Spratlen, James Tobin, and Charles Z. Wilson.

8. *International Economic Association* (Fritz Machlup). As one of the Association's representatives to the International Economic Association, Mr. Machlup gave an oral report on the activities of the latter organization. His written report is printed below.

9. *Other Reports*. Written reports by other Committees and Representatives were summarized by the Secretary and are printed below.

10. *Clearinghouse Proposal* (Vittorio Bonomo). Mr. Bonomo appeared before the Executive Committee to request that it appoint three members of an advisory committee to

his proposed Clearinghouse for Research in Economics. After some discussion, the President appointed a committee consisting of John R. Meyer, Chairman, G. H. Borts, Robert Eisner, and Mark Perlman to consider the Clearinghouse proposal and to submit recommendations to the Executive Committee at its March meeting as to the appropriate role for the Association. The Committee was also asked to consider larger issues of a related nature.

11. *Structure of the Profession.* It was agreed that the Advisory Committee on the Study of the Structure of the Profession be reactivated to consider the problem of analyzing data from the National Registry of the National Science Foundation and to consider the problems discussed in the paper by Allan M. Cartter ("Whither the Market for Academic

Economists?", printed above). The President announced that he was appointing to the Committee Robert Strotz, Chairman, Allen Bonnell, Anne P. Carter, Allan M. Cartter, W. Lee Hansen, and Marshall A. Robinson.

12. *Pareto Project.* The Secretary reported that John Chipman would shortly submit a written report containing recommendations on the Pareto Project. It was agreed that action would be taken by mail ballot when the report is received.

13. *Social Science Research Council.* In separate resolutions, it was VOTED to extend the appreciation of the Association to Paul Webbink and Elbridge Sibley on the occasion of their retirement from the Social Science Research Council and to ask the Secretary to convey its sentiments to them.

The meeting was adjourned at 11:00 p.m.

REPORT OF THE SECRETARY FOR THE YEAR ENDING DECEMBER 31, 1970

Move of Offices. The offices of the Association were moved from Northwestern University, Evanston, Illinois, to Vanderbilt University, Nashville, Tennessee, in mid-September, at which time Rendigs Fels succeeded Harold F. Williamson as Secretary, Treasurer, and Editor of the Papers and Proceedings. A new staff was employed with Charles T. Reich as the Administrative Director.

Annual Meetings. The final report for the 1969 annual meeting of the Allied Social Science Associations, held in New York, N. Y., indicated a total registration of 8,628, including multiple registrations. Approximately 7,000 individuals registered, of whom 64 percent preregistered. The net income from the meeting was \$22,672.98. The Association's share, based on registration figures, was \$15,517.75.

The schedule for future meetings is: 1971, New Orleans, Jung; 1972, Toronto, Royal York; 1973, New York, New York Hilton; 1974, San Francisco, San Francisco Hilton; 1975, Dallas, Dallas Hilton; 1976, Atlantic City, Chalfonte-Haddon Hall

The referendum taken on the time of the annual meetings showed a preference for a date immediately before or after Labor Day, rather than between Christmas and New Year's Day. The vote was 2,405 to 2,044. In accordance with a vote by the Executive Committee, an effort will be made to reschedule the annual meetings for 1976 and earlier years.

Membership. Table 1 shows that the total number of members and subscribers as of December 31, 1970, was 26,173, a net increase of 87 for the year.

TABLE 1—MEMBERS AND SUBSCRIBERS

	Totals 12/31/69	Gain or Loss	Totals 12/31/70
Class of membership:			
Annual.....	16,036	39	16,075
Junior.....	2,308	-231	2,077
Family.....	274	15	289
Complimentary...	151	9	160*
Life.....	276	15	291
Honorary.....	16		16
Total members....	19,061	-153	18,908
Subscribers.....	7,025	240	7,265
Totals.....	26,086	87	26,173

* Includes 93 who did not receive the publications.

Handbook. The 1969 edition of the Handbook was distributed free in February 1970 to all members. In addition, 910 copies have been sold. A comparison of the cost and size with the editions of 1957 and 1965 is included in Table 2. (The 1958 and 1966 editions were supplementary to the complete Handbooks of 1957 and 1965.)

National Register of Scientific and Technical Personnel. Work was completed in Evanston for the 1970 National Register of Scientific and Technical Personnel for the National Science Foundation. The contract between the Association and the NSF has been renewed to allow for preliminary work on the 1972 Register and to provide funds for reactivating the Advisory Committee on the Study of the Structure of the Profession.

Employment Service. The Association continues to sponsor three types of employment service. One is the "Vacancies and Applications" section of the *American Economic Review*. The second is the National Registry for Economists, operated

TABLE 2—PUBLICATION COSTS

Year*	Papers and Proceedings			Handbooks		
	Number of pages	Number of copies	Cost	Number of pages	Number of copies	Cost
1957	754	12,400	\$16,253	548	10,100	\$15,815
1958	677	12,700	15,471	32	9,300	1,434
1959	689	14,000	16,780			
1960	745	14,800	18,914			
1961	675	15,800	19,759			
1962	615	16,000	18,277			
1963	753	17,700	23,440			
1964	698	18,500	23,362	472	16,000	48,626
1965	652	20,000	23,264			
1966	670	22,500	28,405	184	23,000	13,261
1967	741	24,000	32,682			
1968	751	26,500	36,637			
1969	629	27,000	41,263	612	21,900	115,068
1970	533†	28,000	43,031			

* This is the year of publication and pertains to the meeting of the preceding year.

† Small type.

on a year-round basis by the Chicago Placement Service of the Illinois State Employment Service. The third is the employment service provided by the state offices of the U. S. Employment Service at the annual meetings of the Association. The 1969 placement service operation of the Allied Social Science Associations processed 1,800 job openings and 2,100 job applications.

Group Flights. The Association has for a number of years cooperated with a travel agency which has arranged group flights for members attending the annual meetings. In addition, an overseas group flight to the South Pacific has been scheduled for the summer of 1971. Neither the American Economic Association nor any other of the Allied Social Science Associations sponsors or endorses either kind of group flight nor does it accept responsibility for their promotion, sale or operation.

Use of the Mailing List. The continued policy of the Association is to grant requests for the use of the mailing list only if the material appears to be of interest to a substantial proportion of its members. Net income from the sale of the mailing

list during 1970 was \$16,827 compared to \$12,591 for 1969.

A change in the bylaws adopted in 1970 permits nominations to offices in the Association by petition. (Details appear in the minutes of the Annual Meeting, printed above.) The mailing list will be made available at cost to members desiring to circulate such petitions.

Advertising Announcements. There were 179½ pages of paid advertising in the *American Economic Review* and the *Papers and Proceedings* in 1970, compared to 179 pages of paid advertising and 13½ pages of exchange advertising in 1969. (The exchange advertisements were discontinued after the March 1969 issue of the *Review*.) Advertising income was \$50,492 compared to \$50,272 in 1969.

Permission to Reprint and Translate. Official permission to quote from, reprint, or translate and reprint articles from the *American Economic Review* and the *Papers and Proceedings* totaled 202 in 1970 compared to 183 in 1969. Upon receipt of a request to reprint, the publisher-editor is instructed to get the author's permission and to send a copy of his letter of consent

to the Association as a condition for official permission. It is also suggested that a fee of \$150 be paid to the author.

Papers and Proceedings. A comparison of the size and cost of the *Papers and Proceedings* for the period 1957-70 is shown

in Table 2. Despite the apparent reduction in size, the 1970 issue was the largest ever in terms of manuscript pages, making it necessary to use smaller type. In the current (1971) issue, larger type similar to earlier years is being used.

Standing Committees

ADVISORY COMMITTEE ON THE HISTORY OF THE ASSOCIATION

George J. Stigler, *Chairman*
Harold F. Williamson, *Corresponding Secretary*
Joseph Dorfman

CENSUS ADVISORY COMMITTEE

Bert G. Hickman, *Chairman* (1972)
Daniel Creamer (1972)
Donald J. Daly (1971)
Paul A. David (1973)
Solomon Fabricant (1972)
Gary Fromm (1970)
Raymond W. Goldsmith (1972)
Zvi Griliches (1973)
Fred H. Klopstock (1974)
Sherman J. Maisel (1971)
Arthur M. Okun (1972)
Guy H. Orcutt (1970)
Alice Rivlin (1974)
Murray L. Weidenbaum (1971)
Ernest W. Williams (1974)

CENSUS ADVISORY COMMITTEE, TECHNICAL SUBCOMMITTEE ON BUSINESS CYCLE DEVELOPMENTS

Geoffrey H. Moore, *Chairman* (1972)
Donald J. Daly (1971)
Gary Fromm (1971)
Arthur M. Okun (1972)
Beryl W. Sprinkel (1971)
Kenneth B. Williams (1971)

COMMITTEE ON ECONOMIC EDUCATION

G. L. Bach, *Chairman* (1970)
Kenneth E. Boulding (1972)

Marshall Colberg (1972)
Allen C. Kelley (1972)
Henry H. Villard (1973)
Harold F. Williamson (1972)
Rendigs Fels, *Ex Officio*

COMMITTEE ON HONORS AND AWARDS

William H. Nicholls, *Chairman* (1972)
Kermit Gordon (1972)
Richard A. Musgrave (1970)
Douglass C. North (1972)
Anthony D. Scott (1970)
Robert M. Solow (1970)

COMMITTEE ON RESEARCH AND PUBLICATIONS

John Perry Miller, *Chairman* (1971)
Zvi Griliches (1970)
Bert F. Hoselitz (1971)
James N. Morgan (1971)
Albert Rees (1972)
Willard L. Thorp (1972)
Rendigs Fels, *Ex Officio*

ECONOMICS INSTITUTE, POLICY AND ADVISORY BOARD

Henry Rosovsky, *Chairman* (1971)
Albert Fishlow (1973)
Earl O. Heady (1973)
Gustav Ranis (1972)
Joseph J. Spengler (1971)
Jeffrey G. Williamson (1972)

JOINT COMMITTEE WITH THE ASSOCIATION OF AMERICAN LAW SCHOOLS

Gerald M. Meier, *Chairman*

Louis De Alessi
 Harold Demsetz
 Richard A. Musgrave
 Peter O. Steiner
 George J. Stigler

JOINT COUNCIL ON ECONOMIC EDUCATION,
 SPECIAL ADVISORY COMMITTEE

G. L. Bach (1970)
 Rendigs Fels (1971)
 Henry Villard (1972)

Committees Appointed During the Year

COMMITTEE ON ELECTIONS

James S. Worley, *Chairman*
 William O. Thweatt
 Rendigs Fels, *Ex Officio*

COMMITTEE ON FORMATION OF INSTITUTE
 FOR BLACK STUDENTS

Charles L. Schultze
 Arnold C. Harberger

FINANCE COMMITTEE

C. Wells Farnham, *Chairman*
 Corliss D. Anderson
 Milton Friedman
 Harold F. Williamson

ADVISORY COMMITTEE TO THE BLACK CAU-
 CUS

James Tobin, *Chairman*
 Charles L. Schultze
 George H. Borts
 Robert Dorfman
 Marcus Alexis

NOMINATING COMMITTEE

Milton Friedman, *Chairman*
 Irma G. Adelman
 Andrew F. Brimmer
 Hollis B. Chenery
 Albert Fishlow
 Harold M. Somers

COMMITTEE ON THE EDUCATION AND
 TRAINING OF MINORITY GROUP
 ECONOMISTS

Andrew F. Brimmer, *Chairman*
 Frank G. Davis
 Arnold C. Harberger
 Vivian W. Henderson
 Henry Rosovsky
 Charles L. Schultze
 Henry S. Terrell
 Jeffrey G. Williamson
 Wyn F. Owen, *Ex Officio*
 Harold F. Williamson, *Ex Officio*

COMMITTEE TO NOMINATE CANDIDATES
 FOR THE OFFICE OF SECRETARY-
 TREASURER

George J. Stigler, *Chairman*
 William J. Fellner
 Wassily Leontief

Council and Other Representatives

AMERICAN ASSOCIATION FOR THE AD-
 VANCEMENT OF SCIENCE

Joseph A. Pechman (1971)

COMMITTEE ON QUANTITATIVE DATA

Richard A. Easterlin

AMERICAN COUNCIL OF LEARNED SOCIE-
 TIES

Ralph L. Andreano (1970)

CONSORTIUM OF PROFESSIONAL ASSOCIA-
 TIONS FOR THE STUDY OF TEACHER
 IMPROVEMENT PROGRAMS

Henry H. Villard (1972)
 Harold F. Williamson (1970)

INTERNATIONAL ECONOMIC ASSOCIATION

Fritz Machlup (1970)
John T. Dunlop (1973)

NATIONAL ARCHIVES ADVISORY COUNCIL,
GENERAL SERVICES ADMINISTRATION

Harold F. Williamson (1972)

NATIONAL BUREAU OF ECONOMIC RE-
SEARCH

William L. Thorp (1972)

NATIONAL COMMITTEE FOR THE REVISION
OF HANDBOOK II

Ernest Bartell

NATIONAL RESEARCH COUNCIL, DIVISION
OF BEHAVIORAL SCIENCES

Guy H. Orcutt (1971)
Franco Modigliani (1972)
Robert Eisner (1973)

SOCIAL SCIENCE RESEARCH COUNCIL

Zvi Griliches (1970)
James N. Morgan (1971)
Albert Rees (1972)

STANDING COMMITTEE ON TRANSPORTA-
TION AND LONG RANGE PLANNING

Leon N. Moses

*Representatives of the Association on Various Occasions*UNIVERSITY OF AKRON CENTENNIAL CON-
VOCATION

Clinton L. Warne

PROFESSIONAL WOMEN'S CAUCUS

Marion Gillim

INAUGURATIONS

Adolph Gustaf Anderson, Hartwick Col-
lege

Philip Robbins

John W. Kneller, Brooklyn College of The
City University of New York

James C. Cooper

Roy Bryant Shilling, Jr., Hendrix College
David L. Burrough

Robert J. Nossen, Bloomsburg State Col-
lege

Samuel A. Rosenberg

Powell Alexander Fraser, King College

Gordon Ludolf

John Edward Corbally, Jr., Syracuse Uni-
versity

Lionel W. McKenzie

William Courtney Spencer, Western Col-
lege

William J. McKinstry

Frank N. Elliott, Rider College

Max Gideonse

Matthew T. Conlin, Siena College

Morris A. Copeland

John Horton Allen, Centenary College of
Louisiana

David C. Townsend

Donald F. Hornig, Brown University

Nancy Ruggles

Allan Andrew Kuusisto, Hobart & William
Smith Colleges

Hans H. Bernt

Aldrich Kossuth Paul, Upper Iowa Univer-
sity

Calvin D. Siebert

Robert Coldwell Well, University of Mas-
sachusetts at Amherst, Boston and
Worcester

Paul R. Nichols

Gregory H. Adamian, Bentley College

Alice E. Bourneuf

RENDIGS FELS, *Secretary*

REPORT OF THE TREASURER FOR THE YEAR ENDING DECEMBER 31, 1970

A summary comparison of major income and expense items for 1969 and 1970 is given in Table 1 of this report. Detailed results of financial operations for the twelve-month periods ending December 31, 1969, and December 31, 1970, are shown in the Auditors' Reports printed below. The deficit, which was \$6,063 in 1968, rose to \$108,812 in 1969 and \$242,128 in 1970. The expansion of the *Journal of Economic Abstracts* into the *Journal of Economic Literature* was the main reason for the increased deficit in

1969. Rising costs associated with inflation, certain nonrecurring expenses, and capital losses from declining securities markets account for the further rise in 1970.

The operating deficit for 1970 was initially covered by borrowing from the State National Bank of Evanston. The total of loans outstanding reached a maximum of \$175,000. Sale of securities in December reduced outstanding loans to \$75,000 by the end of the year.

RENDIGS FELS, *Treasurer*

TABLE 1—SUMMARY OF INCOME AND EXPENDITURES, 1969–70

	1969	1970
Income		
Dues and activities.....	\$347,715	\$351,361
Investment income.....	25,755	(16,054)*
Total.....	\$373,470	\$335,307
Expenses		
Publications		
Printing and mailing of		
<i>American Economic Review</i>	\$108,636	\$108,579
<i>Journal of Economic Literature</i>	103,221	117,577
<i>Papers and Proceedings</i>	41,263	43,031
Editorial Offices: Salaries and expenditures.	109,764	101,723
Handbook.....		40,817
Other.....	10,250	13,459
Total publication expenses.....	\$373,134	\$425,186
Administration.....	109,148	146,139
Sundry.....		6,110
Total.....	\$482,282	\$577,435
Excess of income over expense.....	\$(108,812)	\$(242,128)

* Interest paid on loans contracted by the Association to cover operating deficits has been deducted from investment income for 1970.

REPORT OF THE FINANCE COMMITTEE

*Executive Committee
American Economic Association
Nashville, Tennessee*

Gentlemen:

The accompanying inventory lists the securities held by the American Economic Association as of December 31, 1970, with the costs and the market values as of that date. The total market value of the securities portfolio at the end of the year was \$493,629. After making adjustments for cash additions and withdrawals, we estimate that the Association's portfolio (not including the income generated by the account) declined in value by slightly less than 4% during 1970. As a basis of comparison, the Dow Jones Industrial Average increased by 4.8% and the New York Stock Exchange Composite Index decreased by 2.5%. While these results may appear somewhat disappointing, we believe that the record is satisfactory, perhaps better than one might have expected, given the fact that the portfolio has a definite—and deliberate—growth orientation and that growth stocks were unusually depressed in the first eight months of the year. In this connection, it should be noted that the account performed appreciably better than the market as a whole during the three prior years (in 1969 the account was down 7% and the DJIA was down 15%; in 1968, the account was up 17% and the DJIA was up 4%; and in 1967 the account was up 19% and the DJIA was up 15%).

It should be remembered that the \$493,629 figure referred to above includes a Special Grant that was made by the Ford Foundation in January of 1969 and subsequently commingled with the Association's

account. As of December 31, 1970, the Associations' portion of the aggregate account was \$334,581, or 67.8%, and the Special Grant represented the remaining \$159,048, or 32.2% of the total. These percentages reflect the \$100,000 withdrawal which was made from the Association's portion of the account on December 22.

During the current year, there were no important changes in the bond portion of the investment account. In the equity portion of the account, a number of changes were made. Issues sold were 550 Royal Dutch, 300 Wells Fargo, 150 Motorola, 500 Public Service of Indiana, 300 Deere & Co., 300 Rex Chainbelt, 250 International Nickel, 600 Archer Daniels Midland, 400 J. Ray McDermott, and 150 Honeywell. Stocks purchased were 200 Continental Can, 300 Moore Corp., 300 Utah Construction, 200 Control Data, 300 Insurance Company of North America, 500 T. I. Corp. (Title Insurance), 100 Merck, 300 Western Bancorporation, and 500 Hoerner Waldorf.

The managers of the investment fund believe that the Association's large equity exposure should be maintained and probably increased as attractive investment opportunities present themselves. The dramatic decline in interest rates, and consequent increase in bond prices during the past six months may well provide a favorable opportunity for transfer of capital from long term bonds to common stocks. This suggests, of course, a belief that the economy is now, or soon to be, in an expansionary stage with increasing corporate profits and stock prices and concern over inflation on the part of investors.

C. WELLS FARNHAM, *Chairman*

TABLE 1—INVENTORY AND APPRAISAL OF SECURITIES AND CASH AS OF 12/31/70

	Par or Shares	Market Value	Cost
CASH AND SHORT-TERM SECURITIES			
CASH.....		1,073	1,073
<i>Corporate and Other</i>			
General Finance Notes.....	6,000	6,000	6,000
TOTAL CASH AND SHORT-TERM SECURITIES.....		7,073	7,073
LONG-TERM BONDS AND PREFERRED STOCKS			
<i>Corporate and Other</i>			
Anheuser Busch.....	10,000	8,400	10,000
Avco.....	20,000	14,200	17,850
Corn Products.....	10,000	8,200	9,950
General Tel. and Elec.....	20,000	16,400	20,530
Marcor, Inc.....	20,000	16,400	16,975
Pet, Inc.....	10,000	7,900	9,900
TOTAL LONG-TERM BONDS AND PREFERRED STOCK.....		71,500	85,205
TOTAL CASH AND FIXED INCOME SECURITIES.....		78,573	86,278
COMMON STOCK AND OTHER CONVERTIBLES			
<i>Utilities</i>			
Central and Southwest.....	300	14,700	2,100
Houston Lighting and Power.....	300	13,800	4,992
		28,500	7,092
<i>Financial</i>			
Banco Popular Puerto Rico.....	385	14,630	17,675
First Bank System.....	400	24,400	12,690
INA Corp.....	500	18,500	17,500
Western Bancorporation.....	300	11,400	10,022
		68,930	57,887
<i>Foods and Containers</i>			
Continental Can.....	300	11,700	13,642
<i>Papers</i>			
Hoerner Waldorf.....	500	12,000	12,107
<i>Housing Related</i>			
Boise Cascade.....	255	11,730	11,445
Sherwin Williams.....	300	14,100	16,771
T I Corp.....	500	15,000	12,924
		40,830	41,140
<i>Mining and Metals</i>			
McIntyre Porcupine Mines.....	100	14,200	4,818
Utah Construction and Mining.....	300	16,800	15,132
		31,000	19,950
<i>Oil and Gas</i>			
Atlantic Richfield.....	200	13,200	9,625
Pennzoil United.....	200	9,600	15,293
		22,800	24,918
<i>Drugs and Medical</i>			
Abbott Laboratories.....	300	22,800	6,135
American Hospital Supply.....	400	14,000	15,754
Merck.....	100	9,900	9,025
		46,700	30,914
<i>Electrical Products</i>			
Westinghouse Electric.....	200	13,400	11,615
Zenith Radio.....	300	11,100	1,975
		24,500	13,590
<i>Computers</i>			
Control Data.....	200	10,200	8,495
International Business Machines.....	122	38,796	10,169
		48,996	18,664

REPORT OF THE FINANCE COMMITTEE

489

TABLE 1—(Continued)

	Par or Shares	Market Value	Cost
<i>Other Office Equipment</i>			
Moore Corp.....	300	11,400	10,950
Xerox.....	150	13,050	5,305
		24,450	16,255
<i>Miscellaneous</i>			
Disney (Walt) Productions.....	300	42,300	5,241
Donnelley (R. R.) & Sons.....	650	12,350	14,304
		54,650	19,545
TOTAL COMMON STOCKS AND OTHER CONVS.....		415,056	275,704
TOTAL SECURITIES AND CASH.....		493,629	367,982

AUDITORS' REPORT FOR 1970

To the Executive Committee of

The American Economic Association:

We have examined the balance sheet of THE AMERICAN ECONOMIC ASSOCIATION (a District of Columbia corporation, not for profit) as of December 31, 1970, and the related statements of revenues and expenses and changes in restricted fund balances for the year then ended. Our examination was made in accordance with generally accepted auditing standards, and accordingly included such tests of the accounting records and such other auditing procedures as we considered necessary in the circumstances.

In our opinion, the accompanying balance sheet and statements of revenues and expenses and changes in restricted fund balances present fairly the assets, liabilities and fund balances of the American Economic Association as of December 31, 1970, and its revenues, expenses and changes in fund balances for the year then ended, in conformity with generally accepted accounting principles applied on a basis consistent with that of the preceding year.

ARTHUR ANDERSEN & Co.

Memphis, Tennessee

February 12, 1971.

THE AMERICAN ECONOMIC ASSOCIATION
BALANCE SHEET—DECEMBER 31, 1970

ASSETS		LIABILITIES AND FUND BALANCES		
CASH.....		NOTES PAYABLE TO BANK (Note 3).....		\$ 75,000
INVESTMENTS (Note 1):		ACCOUNTS PAYABLE AND ACCRUED LIABILITIES .		88,266
Corporate bonds and obligations, at market		DEFERRED INCOME:		
value (cost of \$130,205).....	\$116,500	Life membership dues.....	\$ 37,400	
Corporate stocks, at market value (cost of	415,056	Prepaid membership dues.....	28,637	
\$275,187).....		Prepaid subscriptions.....	51,335	117,372
RECEIVABLES, less allowance for doubtful accounts		DEFERRED INCREMENT IN MARKET VALUE OF		
of \$113.....		INVESTMENTS (Note 1).....		50,211
PREPAID EXPENSES.....		RESTRICTED FUND BALANCES.....		214,825
OFFICE FURNITURE AND EQUIPMENT, at cost, less		GENERAL FUND:		
accumulated depreciation of \$1,213.....		Balance, December 31, 1969.....	\$305,937	
		Less—Excess of expenses over revenues	(242,128)	63,809
		for year ended December 31, 1970		<u>\$609,483</u>

The accompanying notes to financial statements are an integral part of this balance sheet.

AMERICAN ECONOMIC ASSOCIATION

THE AMERICAN ECONOMIC ASSOCIATION
STATEMENT OF CHANGES IN RESTRICTED FUND BALANCES
FOR THE YEAR ENDED DECEMBER 31, 1970

	Balance Dec. 31, 1969	Re- ceipts	Disburse- ments	Alloca- tion of Invest- ment Losses (Note 2)	Balance Dec. 31, 1970
The Ford Foundation grants for—					
Translation of foreign economic publications and survey of foreign economic research.....	\$ 13,400	\$ —	\$ (9,273)	\$ —	\$ 4,127
Economics Institute's orientation program for foreign graduate students of economics.....	164,287	39,234	—	(491)	203,030
Funds reserved by the Association for publication of revised editions of <i>Graduate Study in Economics</i> , a guide originally published with funds from a Ford Foundation grant.....	3,425	99	(284)	—	3,240
The Asia Foundation grant for Asian economist membership dues to the American Economic Association and related travel expenses.....	1,342	2,500	(2,000)	—	1,842
The National Science Foundation grant for visiting scientist program.....	3,425	7,036	(8,061)	—	2,400
Sundry.....	327	—	(141)	—	186
	<u>\$186,206</u>	<u>\$48,869</u>	<u>\$(19,759)</u>	<u>\$(491)</u>	<u>\$214,825</u>

The accompanying notes to financial statements are an integral part of this statement.

THE AMERICAN ECONOMIC ASSOCIATION
STATEMENT OF REVENUES AND EXPENSES
FOR THE YEAR ENDED DECEMBER 31, 1970

REVENUES FROM DUES AND ACTIVITIES:				
Membership dues.....				\$173,748
Subscriptions.....				70,235
Advertising.....				56,492
Sale of copies, republications and handbooks.....				11,739
Sale of mailing list.....				16,827
Annual meeting.....				17,597
Sundry.....				4,723
				<u>\$351,361</u>
INVESTMENT LOSSES:				
Corporate bonds and obligations—				
Interest.....	\$ 9,685			
Less—Decline in market value (Note 1).....	(10,257)		\$ (572)	
Corporate stocks—				
Cash dividends.....	\$ 10,219			
Less—Decline in market value recognized (Note 1).....	(18,756)		(8,537)	
Investment losses allocated to a restricted fund (Note 2).....			491	(8,618)
Net revenue.....				<u>\$342,743</u>
EXPENDITURES:				
Publication (Exhibit I).....	\$425,186			
Administrative (Exhibit I).....	146,139			
Interest.....	7,436			
Sundry.....	6,110			584,871
EXPENSES IN EXCESS OF REVENUES.....				<u>\$(242,128)</u>

The accompanying notes to financial statements and Exhibit I are an integral part of this statement.

THE AMERICAN ECONOMIC ASSOCIATION
NOTES TO FINANCIAL STATEMENTS
DECEMBER 31, 1970

- (1) The Association accounts for investments on a market value basis. Under this method, the change in market value of corporate stocks, after adjusting for an inflation factor, is recognized in income over a three-year period. The change in market value of corporate bonds and obligations, after adjusting for an inflation factor, is reflected in income currently.
- (2) In 1968, the Association entered into an agreement with the University of Colorado relating to the Ford Foundation grant for the Economics Institute which provides, among other things, that the Association invest a portion of the funds received and allocate any income and market value adjustments therefrom to the restricted fund. In accordance with this agreement, net investment losses of \$491 were allocated to the restricted fund for the year ended December 31, 1970.
- (3) The Association has pledged certain investment securities as collateral for its notes payable to a bank. At December 31, 1970, these securities had a market value of approximately \$276,000 and a cost of approximately \$146,000.

EXHIBIT I—THE AMERICAN ECONOMIC ASSOCIATION
STATEMENT OF PUBLICATION AND ADMINISTRATIVE EXPENSES
FOR THE YEAR ENDED DECEMBER 31, 1970

PUBLICATION EXPENSES:

Printing and mailing of—	
<i>American Economic Review</i>	\$108,579
<i>Journal of Economic Literature</i>	117,577
<i>Papers and Proceedings</i>	43,031
Handbook.....	40,817
Addressing expenses.....	13,459
Editorial and office salaries.....	66,833
Payments to contributors.....	2,513
Editorial supplies and expense.....	32,377
Total publication expenses.....	<u>\$425,186</u>

ADMINISTRATIVE EXPENSES:

Salaries.....	\$ 83,455
Rent.....	2,819
Postage.....	2,805
Stationery, printing and supplies.....	5,521
Insurance.....	904
Annual meeting expenses.....	12,389
Pension annuity payments.....	3,949
Payroll taxes.....	3,049
Depreciation (straight-line method).....	822
Telephone and telegraph.....	1,510
Committee on Research & Publications.....	2,499
Committee expenses for Caucus of Black Economists.....	10,000
Executive Committee expenses.....	2,093
Other committees' expenses.....	5,184
President and president-elect expenses.....	2,000
Sundry.....	10,516
Expenses allocated to restricted funds.....	(3,376)
Total administrative expenses.....	<u>\$146,139</u>

AUDITORS' REPORT FOR 1969

*To the Executive Committee of
American Economic Association:*

We have examined the statement of assets and liabilities of American Economic Association (a District of Columbia corporation, not for profit) as of December 31, 1969, and the related statements of changes in restricted fund balances and income and expenses for the year then ended. Our examination was made in accordance with generally accepted auditing standards, and accordingly included such tests of the accounting records and such

other auditing procedures as we considered necessary in the circumstances.

In our opinion, the accompanying statements present fairly the assets and liabilities of American Economic Association as of December 31, 1969, and the changes in restricted fund balances and income and expenses for the year then ended, in conformity with generally accepted accounting principles applied on a basis consistent with that of the preceding year.

ARTHUR ANDERSEN & Co.

Chicago, Illinois,
January 28, 1970.

EXHIBIT 1—AMERICAN ECONOMIC ASSOCIATION, STATEMENT OF ASSETS AND LIABILITIES, DECEMBER 31, 1969

ASSETS		LIABILITIES	
CASH.....		ACCOUNTS PAYABLE.....	\$ 57,547
INVESTMENTS (Note 1):		DEFERRED INCOME:	
Corporate bonds and obligations at		Life membership dues.....	\$ 33,100
market value (cost of \$148,086).....	\$113,800	Prepaid membership dues.....	27,322
Corporate stocks, at market value		Prepaid subscriptions.....	36,942
(cost of \$310,458).....	500,340		
		DEFERRED INCREMENT IN MARKET VALUE OF	
RECEIVABLES, less allowance for doubtful		INVESTMENTS (Note 1).....	56,602
accounts of \$1,700.....		RESTRICTED FUND BALANCES (Exhibit 2).....	186,206
PREPAID EXPENSES.....	35,383	GENERAL FUND:	
OFFICE FURNITURE AND EQUIPMENT, at cost, less	950	Balance, December 31, 1968.....	\$414,749
accumulated depreciation of \$3,690.....	3,828	Less—Excess of expenses over income for the	(108,812)
		year ended December 31, 1969 (Exhibit 3)...	
			305,937
			<u>\$703,656</u>

The accompanying notes to financial statements are an integral part of this statement.

AMERICAN ECONOMIC ASSOCIATION

EXHIBIT 2—AMERICAN ECONOMIC ASSOCIATION, STATEMENT OF CHANGES IN RESTRICTED FUND BALANCES
FOR THE YEAR ENDED DECEMBER 31, 1969

	Balance Dec. 31, 1968	Receipts	Disburse- ments	Allocation of Investment Losses (Note 2)	Balance Dec. 31, 1969
The Ford Foundation grants for—					
Translation of foreign economic publica- tions and survey of foreign economic research.....	\$14,987	\$ —	\$ (1,587)	\$ —	\$ 13,400
Foreign student screening and selection..	11,846	631	(9,052)	—	3,425
Economics Institute's orientation pro- gram for foreign graduate students of economics.....	—	206,000	(36,000)	(5,713)	164,287
United States Government grant for foreign student screening and selection.....	4,470	—	(4,470)	—	—
The Asia Foundation grant for Asian econo- mist membership dues to the American Economic Association and related travel expenses.....	1,391	—	(49)	—	1,342
The National Science Foundation grant for visiting scientist program.....	3,748	6,345	(6,668)	—	3,425
Sundry grants.....	227	200	(100)	—	327
	<u>\$36,669</u>	<u>\$213,176</u>	<u>\$(57,926)</u>	<u>\$(5,713)</u>	<u>\$186,206</u>

The accompanying notes to financial statements are an integral part of this statement.

EXHIBIT 3—AMERICAN ECONOMIC ASSOCIATION, STATEMENT OF INCOME AND EXPENSES
FOR THE YEAR ENDED DECEMBER 31, 1969

INCOME FROM DUES AND ACTIVITIES:			
Membership dues.....		\$ 163,194	
Subscriptions.....		75,268	
Advertising.....		50,272	
Sale of copies, republications and handbooks.....		10,631	
Sale of mailing list.....		12,591	
Annual meeting.....		32,440	
Sundry.....		3,319	
		<u>\$ 347,715</u>	
INVESTMENT INCOME:			
Corporate bonds and obligations—			
Interest.....	\$15,548		
Less- Decline in market value (Note 1).....	35,585	\$(20,037)	
Corporate stocks—			
Cash dividends.....	\$12,422		
Increment in market value recognized (Note 1).....	27,657	40,079	
		<u>5,713</u>	
Investment losses allocated to a restricted fund (Note 2).....			25,755
			<u>\$ 373,470</u>
EXPENSES:			
Publication (Exhibit 4).....		\$362,884	
Administrative (Exhibit 4).....		109,148	
Grant to Yale University for <i>Index of Economic Journals</i>		10,250	482,282
			<u>\$ (108,812)</u>
Excess of expenses over income for the year.....			

The accompanying notes to financial statements and Exhibit 4 are an integral part of this statement.

AUDITORS' REPORT FOR 1969

497

EXHIBIT 4—AMERICAN ECONOMIC ASSOCIATION, STATEMENT OF PUBLICATION AND ADMINISTRATIVE EXPENSES FOR THE YEAR ENDED DECEMBER 31, 1969

PUBLICATION EXPENSES:

Printing and mailing of—	
<i>American Economic Review</i>	\$108,636
<i>Journal of Economic Literature</i>	103,221
<i>Papers and Proceedings</i>	41,263
<i>Handbook</i>	20,775
Addressing expenses.....	10,801
Editorial and office salaries.....	61,885
Payments to contributors.....	4,548
Editorial supplies and expense.....	11,755
Total publication expenses.....	<u>\$362,884</u>

ADMINISTRATIVE EXPENSES:

Secretary's salary, less \$2,250 allocated to restricted funds.....	\$ 10,593
Office salaries.....	47,029
Office machine rentals.....	203
Postage.....	5,493
Stationery, printing and supplies.....	2,933
Insurance.....	788
Annual meeting expenses.....	14,620
Executive Committee expenses.....	2,351
Other committees' expenses.....	2,239
Archives.....	60
Pension annuity payments.....	3,537
Payroll taxes.....	2,784
Depreciation of furniture and equipment.....	563
Telephone and telegraph.....	1,553
Committee on Research & Publications.....	125
President and president-elect expenses.....	1,607
Sundry.....	13,582
Expenses allocated to restricted funds.....	(912)
Total administrative expenses.....	<u>\$109,148</u>

NOTES TO FINANCIAL STATEMENTS
DECEMBER 31, 1969

- (1) The Association accounts for investments on a market value basis. Under this method, the change in market value of corporate stocks, after adjusting for an inflation factor, is recognized in income over a three-year period. The change in market value of corporate bonds and obligations, after adjusting for an inflation factor, is reflected in income currently.
- (2) In 1968, the Association entered into an agreement with the University of Colorado relating to the Ford Foundation grant for the Economics Institute which provides, among other things, that the Association invest a portion of the funds received and allocate any income and market value adjustments therefrom to the restricted fund. In accordance with this agreement, net investment losses of \$5,713 were allocated to the restricted fund for the year ended December 31, 1969.

AMERICAN ECONOMIC REVIEW

REPORT OF THE MANAGING EDITOR FOR THE YEAR ENDING DECEMBER 1970

The number of manuscripts submitted during 1970 was 879, 121 more than the preceding year and more than double the number submitted in 1965. While the rate of growth has slowed somewhat from earlier years, I expect the number to reach 1,000 in 1971. These figures are summarized in Table 1, which shows the number of manuscripts submitted and published since 1951. While the acceptance rate is lower than last year's 16 percent, this average figure should in no way be construed as the probability that a good manuscript will be published. Many of the papers sent in represent preliminary research reports which are then reworked and revised by the authors, with ultimate publication in a learned journal.

The distribution of submitted and published manuscripts by subject matter is shown in Table 2. The most popular fields this year are micro theory, international

TABLE 1—MANUSCRIPTS SUBMITTED AND PUBLISHED, 1951–1970

Year	Submitted	Published	Ratio of Published to Submitted
1951	222	49	.22
1952	190	47	.25
1953	234	51	.22
1954	231	47	.20
1955	245	41	.17
1956	242	48	.20
1957	215	40	.19
1958	242	46	.19
1959	279	48	.17
1960	276	46	.17
1961	305	47	.15
1962	273	46	.17
1963	329	46	.14
1964	431	67	.16
1965	420	59	.14
1966	451	62	.14
1967	534	94	.18
1968	637	93	.15
1969	758	121	.16
1970	879	120	.14

economics, economic growth, and welfare economics. Almost all fields are represented in the *Review*, although not in ev-

TABLE 2—SUBJECT-MATTER DISTRIBUTION OF SUBMITTED AND PUBLISHED MANUSCRIPTS IN 1970

	Submitted	Published
General economics—theory; history; systems		
General economics and general equilibrium theory	26	1
Microeconomic theory	177	15
Macroeconomic theory	47	2
Welfare theory	79	18
Economic history; history of thought; methodology	17	3
Economic systems	6	—
Economic growth; development; planning; fluctuations	94	15
Economic statistics	18	9
Monetary and financial theory and institutions	83	8
Fiscal policy and public finance	24	—
International economics	90	20
Administration; business finance	14	4
Industrial organization	27	2
Agriculture; natural resources	27	3
Manpower; labor; population	78	8
Welfare programs; consumer economics; urban and regional economics	50	11
Other	22	1

TABLE 3—SUMMARY OF CONTENTS, 1969-70

	1969		1970	
	Number	Pages	Number	Pages
Articles.....	47	579	52	657
Communications, including notes, comments, replies..	67	303	68	334
Special articles.....	7	55	—	—
Total.....	121		120	
Dissertations.....		46		75
Announcements and Notes.....		47		36
		1,030		1,102

ery year. There has been a growth in the number of publishable papers in the fields of pure and applied welfare economics, and I think this reflects the research activities in the profession.

Table 3 summarizes the contents of the four regular issues of the *Review* during 1970, comparing it with the previous year. Without design, the average size of articles has stabilized at 12.5 journal pages and of communications at 5 journal pages. No doubt the average size of articles will have to be reduced in the future to make room for increased numbers of submissions. The total size of the journal in 1970 (1,102 net pages) exceeded my plans, in part because of the unusually large annual list of doctoral dissertations which now appears in the December issue. It is of interest to note that almost 1,000 Ph.D. dissertations in economics and business administration are listed together with brief 75-word abstracts. Many of these will show up as journal articles in the future.

Because of the large number of submissions, I introduced a new method of handling manuscripts submitted to the *Review*. I have made use of the services of younger members of the profession who act as editorial consultants. They read manuscripts and aid me in determining their suitability for detailed refereeing. This screening task is feasible because many submitted papers are either in the form of preliminary research reports or represent work which is already available in some printed medium. In no case, however, is a paper published without formal refereeing. Because each editorial consultant will screen from twenty to forty manuscripts at once, they are paid a stipend.

The 1970 costs of printing the *Review* are shown in Table 4 and shown for other years in Table 5. The gross number of pages exceeds the net because of advertising and index pages. Printing costs rose in 1970 over 1969 because of greater number

TABLE 4—COPIES PRINTED, SIZE AND COST OF PRINTING AND MAILING IN 1970

	Printed	Pages		Cost		
		Net	Gross	Issue	Reprints	Total
March.....	28,000	264	340	\$ 28,338.02	\$257.20	\$ 28,595.22
June.....	28,000	266	304	25,544.57	283.85	25,828.52
September.....	27,970	258	308	25,644.78	54.69	25,699.47
December*.....	27,000	314	384	31,500.00	200.00	31,700.00
	110,970	1,102	1,336	\$111,027.47	\$795.74	\$111,823.21

* Estimate.

TABLE 5—ACTUAL AND BUDGETED EXPENDITURES 1965-1971

	Printing and Mailing	Payments to Contributors	Office Expenses	Total
1965	\$ 61,606	\$1,598	\$22,486	\$ 85,691
1966	74,277	1,873	24,844	100,996
1967	82,840	2,255	27,114	112,209
1968	92,948	2,140	29,155	124,243
1969	97,183	600	26,765	124,548
1970	(Actual) 111,823	—	36,336	148,159
1970	(Budget) 104,701	—	35,474	140,175
1971	(Budget) \$129,534	—	42,848	172,382

of copies, the larger number of pages, and higher printing charges. However, the number of pages printed will not be allowed to increase. Further, the number of copies per issue will be held at 27,000 to keep the inventory backlog at a lower level.

Office expenses have also increased in 1970 over 1969, as shown in Table 5. Part of the increase represents higher salaries, and part the cost of running the editorial consultant program for six months. I expect office expenses to rise again in 1971 because of higher salaries and increases in postage and supplies. Moreover, the editorial consultant program will run for a full year in 1971, thus increasing the budgeted cost.

Four members of the Board of Editors complete their terms at this time: Gary Becker, Irwin Friend, Lloyd Ulman, and Warren Smith. I wish to thank them for their cooperation in the face of strenuous demands on their time. Their work has served to improve the quality of the articles printed.

I am submitting to the Executive Committee the names of the following nominees to the Board of Editors, each to serve three year terms: Phillip Cagan, Daniel McFadden, Herbert Mohring, and Edmund S. Phelps.

I should like to express my thanks to members of the Board of Editors whose terms are continuing: Barbara Bergmann, Jagdish Bhagwati, Gregory Chow, Carl Christ, Charles Ferguson, Harry Johnson, Alvin Marty, Marc Nerlove, Warren Nutter, and Vernon Smith. They work patiently and with great devotion and conscience to preserve the quality of the journal.

The following individuals assisted me as editorial consultants: Barry Chiswick, Laurits Christensen, Mark Daniels, George Eads, Henry Grabowski, E. Philip Howrey, and Lawrence Lau.

I have received valuable proofreading and mathematical assistance from Martin Ringo and William Holahan. Finally, I should like to thank Mrs. Wilma St. John for her fine work as assistant editor.

In addition to the members of the Board and the editorial consultants, I have sought and received the assistance of a large number of economists during the course of the year. I wish to thank them for their cooperation and recognition of high standards in reading and evaluating manuscripts. They have eased the heavy work load that would have otherwise fallen on the Board of Editors. The following have assisted in this way:

H. Aaron
I. Adelman
M. A. Adelman

D. Aigner
G. A. Akerlof
P. Albin

A. Alchian
P. Allen
E. Ames

K. J. Arrow
O. Ashenfelter
H. Averch

M. Bailey	G. Crouse	J. Hanson	N. Leff
J. S. Bain	W. Culbertson	J. Harris	H. Leibenstein
B. Balassa	J. Cunyngnam	R. Haveman	A. Leijonhufvud
R. E. Baldwin	R. M. Cyert	G. A. Hay	D. Levhari
W. J. Baumol	M. R. Daniels	P. Hendershott	H. Levin
M. J. Beckmann	E. J. Davis	C. Hildreth	H. S. Levine
S. Besen	O. Davis	J. Hirshleifer	H. H. Liebhaufsky
C. W. Bischoff	L. DeAlessi	J. Hixson	C. Lloyd
A. I. Bloomfield	F. deLeeuw	H. Hochman	M. C. Lovell
M. Blume	M. E. DePrano	R. F. Hoffman	R. E. Lucas, Jr.
R. Boddy	R. K. Diwan	R. Holbrook	R. A. McCain
V. Bonomo	F. T. Dolbear	R. A. Holmes	J. J. McCall
T. E. Borcherding	R. Dorfman	E. S. Howle	J. McKie
S. Bowles	D. J. Dugan	E. P. Howrey	R. McKinnon
D. Bradford	G. Eads	F. S. T. Hsiao	S. P. Magee
W. M. Branson	R. Eckaus	C. G. Hufbauer	B. Malkiel
G. F. Break	R. Ehrenberg	S. Hymer	E. Mansfield
F. P. R. Brechling	R. Eisner	D. Jaffee	J. Margolis
W. Breit	A. Faden	E. James	J. Markham
H. Brems	E. Fama	M. C. Jensen	D. D. Martin
R. Britto	J. Fei	L. Johnson	B. F. Massell
M. Bronfenbrenner	M. Feldstein	B. F. Johnston	T. Mayer
G. Brown	P. Feldstein	D. W. Jorgenson	D. Meiselman
J. Brown	T. J. Finn	M. Kafoglis	A. H. Meltzer
E. Brubaker	A. Fisher	E. Kalachek	J. R. Melvin
J. Buchanan	J. Flanders	D. R. Kamerschen	D. Mermelstein
E. Burmeister	J. E. Floyd	M. Kamien	C. Metcalf
O. Burt	D. Foley	J. H. Kareken	P. Meyer
P. Cagan	K. A. Fox	G. S. Kaufman	R. Meyer
C. D. Campbell	A. Friedlaender	R. Kavesh	M. Miller
W. T. Carleton	H. Galper	D. Keesing	N. Miller
W. Chang	F. Gehrels	A. K. Klevorick	R. L. Miller
A. Charnes	W. E. Gibson	B. P. Klotz	E. S. Mills
S. Cheung	C. Goetz	J. Kmenta	J. Mincer
J. S. Chipman	R. Goldfarb	A. V. Kneese	E. J. Mishan
B. R. Chiswick	S. M. Goldfeld	I. B. Kravis	H. Mohring
V. S. Chitre	H. I. Grossman	M. E. Kreinin	M. Montias
P. B. Clark	M. A. Grove	A. Krueger	S. Morley
R. Clower	J. M. Guttentag	M. Kurz	C. T. Morris
R. Coen	J. Gwartney	M. Ladenson	D. T. Mortensen
R. Cooper	J. Hadar	K. J. Lancaster	J. Mossin
T. J. Courchene	E. T. Haefele	M. Landsberger	D. Mueller
J. C. Cragg	R. Hall	L. J. Lau	R. Muth
M. A. Crew	M. Hamburger	L. B. Lave	M. Myers
J. B. Crockett	B. Hansen	T. H. Lee	E. Nadel
J. G. Cross	W. L. Hansen	L. Lefebvre	M. I. Nadiri

K. Nagatani	R. Radner	D. Schwartzman	W. S. Vickrey
S. Naya	J. Rakowski	T. Scitovsky	P. E. Vincent
P. Neher	L. Rapping	A. D. Scott	M. L. Wachter
R. Nelson	R. Rasche	G. W. Scully	F. V. Walker
J. Newhouse	M. W. Reder	W. F. Sharpe	H. W. Watts
D. Nichols	A. Rees	M. Shubik	R. Waud
Y. Niho	S. Reiter	E. Smolensky	J. Weicher
W. Nordhaus	S. Resnick	D. Smyth	R. L. Weil
J. H. Noren	E. P. Reubens	W. Snyder	R. Weintraub
H. O. Nourse	G. Rich	L. Solmon	L. J. Weiss
J. B. Nugent	R. J. Rohr	R. M. Solow	F. R. Welch
W. Oakland	H. Rose	C. M. Sprenkle	S. H. Wellisz
W. Oi	S. Rosen	R. M. Starr	F. M. Westfield
B. Okner	R. N. Rosett	J. L. Stein	A. Whinston
E. O. Olsen	J. Rothenberg	R. M. Stern	T. M. Whitin
J. Olson	S. Rottenberg	G. Stevens	M. Whitman
M. Olson	R. J. Ruffin	J. E. Stiglitz	H. F. Williamson, Jr.
D. Orr	V. Ruttan	B. Stigum	O. Williamson
M. V. Pauly	H. E. Ryder, Jr.	J. A. Stockfish	R. Wilson
P. Pelikan	A. W. Sametz	R. Sutch	D. M. Winch
S. Peltzman	R. Saposnik	A. H. H. Tan	S. Winter
R. G. Penner	R. Sato	L. D. Taylor	J. H. Wood
G. Perry	T. Saving	H. Theil	D. A. Worcester
C. Phelps	J. Schaefer	E. Thorbecke	W. P. Yohe
C. Plott	F. M. Scherer	A. W. Throop	P. A. Yotopoulos
C. G. Plourde	D. Schilling	L. Thurow	E. S. Yudin
F. Puffer	R. Schmalensee	M. P. Todaro	E. E. Zajac
R. E. Quandt	T. P. Schultz	S. C. Tsiang	P. Zarembka
J. Quirk	M. Schupack	D. L. Tuttle	E. Zupnick
T. Rader	N. L. Schwartz	J. Vanek	

GEORGE H. BORTS, *Managing Editor*

JOURNAL OF ECONOMIC LITERATURE

REPORT OF THE MANAGING EDITOR FOR YEAR ENDING DECEMBER 1970

The 1970 volume of *JEL* is numbered Volume VIII; you may recall that we preserved the numerical succession from the *Journal of Economic Abstracts* in order to protect our second class mailing privileges. We are now publishing on schedule. It contains 1,538 numbered pages plus eight unnumbered pages and four sets of covers. Circulation of issues in this volume averaged over 27,000.

The Contents

There were four general survey articles:

Jacob Mincer, "The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach"

Bruce F. Johnston, "Agriculture and Structural Transformation in Developing Countries: A Survey of Research"

Martin Bronfenbrenner, "Radical Economics in America, 1970"

M. Ishaq Nadiri, "Some Approaches to the Theory and Measurement of Total Factor Productivity"

In addition, there were two interpretative (sometimes called, "creative curmudgeon") articles clearly expressing normative views:

Martin Shubik, "A Curmudgeon's Guide to Microeconomics"

Warren L. Smith, "On Some Current Issues in Monetary Economics: An Interpretation"

We also introduced two bibliographic articles:

Andrew F. Brimmer and Harriet Harper,

"Economists' Perception of Minority Economics: An Interpretation"

Richard W. Lindholm, "The Value Added Tax: A Short Review of the Literature"

Besides these major articles there were six review articles, several communications, and a classified listing of textbooks. There were "normal" length reviews of 177 books, 1,145 annotations of books, plus annotations of 18 "new" journals, tables of contents listings from over 700 journal issues (we list about 250 journals, including 71 foreign language journals), classifications of over 4,700 articles (with about 8,200 total listings), and about 1,025 abstracts of articles.

The number of journals which we list fully or partially has been increasing: first, as a consequence of some journals on the list (inherited from Professor Smithies) sending in back issues and second, because of requests for inclusion from new economics journals and journals only partially devoted to economics topics.

In order to reduce burgeoning costs, we are contemplating a reduction of journal coverage. One idea is to cut back on the listing of articles peripheral to economics from journals with broader coverage, although we could continue to list all their economic articles. Also, while we will continue to list complete tables of contents for foreign language journals, we propose to index only articles in English or articles with an English summary. Since the *Economic Journal* is discontinuing its listing, Mr. Michael Kaser (St. Antony's College, Oxford) has agreed to translate titles

from the Slavic language journals. Again, we will list the tables of contents of these journals but not index them, unless there are English summaries.

We have reduced the length of abstracts from 300 to 100 words. This "rule" became effective during the past year. We have, as might be expected, some carry-over from abstracts submitted or prepared prior to the ruling. However, we will be enforcing the 100-word limit rigorously in the near future. In the main we have not had problems with this limit although there are those who feel that 100 words are "too few." On the other hand, the cost of printing the abstracts is far from small. We are also in the process of reviewing the lists of journals with which we have abstracting agreements. We are trying to strike a compromise between a selection of the "best known" American journals and the better known foreign journals (generally inaccessible to members of the Association).

Editorial Board and Editorial Consultants

During the year Professors Alexander Gershenkron and D. Gale Johnson joined the Board of Editors—replacing Professors Richard A. Easterlin and Karl A. Fox. Professors Peter Kenen and Marc Nerlove will continue to serve through 1971. I shall, during the course of this meeting, request replacements for Professors David I. Fand and Arnold C. Harberger whose terms expire on 31 December 1970. I would also like to add one additional member now and in each of the next two years to eventually have nine members of the Board of Editors rather than six.

All the members of the 1970 Board of Editors were extremely helpful: may I register here my personal gratitude to each of them.

In addition, I have consulted the following (to whom, too, my personal thanks are given), who helped by reading

manuscripts or by giving me advice concerning specific editorial matters:

Irma Adelman	Robert Dorfman
Sanford Berg	Jack Ochs
Abram Bergson	Fritz Machlup
Kenneth Boulding	Alan Peacock
Martin Bronfen-	Alan Prest
brenner	Andreas Papandreou
Murray Brown	Richard Ruggles
Janet Chapman	Ryuzo Sato
A. W. Coats	Theodore W. Schultz
Richard Easterlin	Erik Thorbecke
John G. Gurley	Jerome Wells
Otto Davis	Marina Whitman

Contents for 1971

Survey articles have been commissioned:

- J. E. Morton, "A Student's Guide to Statistical Sources"
- Phyllis Wallace, "Black-White Economic Differentials"
- Edward Budd, "American Income Distribution"
- Thaddeus Spratlen, "The Economics of the Black Ghetto"

"Creative curmudgeon" (or interpretative) articles have been promised by:

- Ezra Mishan (public goods, some theoretical aspects)
- Arnold Harberger (benefit/cost theory)
- Paul Samuelson (the Marxian transformation problem)
- Wilfred Prest (fiscal federalism and revenue sharing in Australia, the United States, and Canada)
- A. W. Coats (the contents of economics journals over time)
- A. Bajt (business cycles in Eastern European economies)

We plan to publish a list of addresses and subscription rates of all economics journals in the September 1971 issue. This step was requested by the journal editors in December 1968 and will be an annual or a biannual activity.

MARK PERLMAN, *Managing Editor*

REPORT OF THE COMMITTEE ON ECONOMIC EDUCATION

During 1970, the main activities of the Association's Committee on Economic Education were the following:

1. A "National Question Bank," for the use of all economists teaching at the undergraduate level, has been established by the Joint Committee on Economic Education, 1212 Avenue of the Americas, New York, New York 10036. This "Question Bank" will include, we hope, a large number of questions covering the various areas of elementary economics, and possibly ranging into more advanced undergraduate courses, collected by the JCEE in collaboration with this Committee or submitted voluntarily by teachers throughout the country. Only questions which have been pretested and on which there is some information as to reliability and performance will be included. Questions will be made available free, or at cost if quantities are desired, by the JCEE. To help establish this Question Bank, a national competition, with a first prize of \$1000 for the best questions, has been established by the JCEE; information was provided in the *American Economic Review* for March 1970.

2. The new *Journal of Economic Education*, published by the Joint Council on Economic Education with the cooperation of this Committee, is firmly established. Its third issue has been published (on a semiannual basis) and paid subscriptions total 3,000. The purpose of this journal is to provide a channel of communication for research work and related publishable material with reference to economic education, primarily at the undergraduate level in colleges and universities, but including junior colleges and, to some extent, the high schools. The flow of manu-

scripts to the journal and the growth of the subscription list have been more rapid than we expected, and the response to the *Journal* has been extremely favorable.

3. Experimental "pilot projects" in teaching basic economics have been established by the Joint Council, with the cooperation of this Committee, at three universities, and additional projects are under consideration. Each project involves an experimental approach to the teaching of the basic course in economics, with careful evaluation of results, in the expectation that the experiment and the results will be made widely available to the profession, especially for the benefit of colleges which find it difficult to undertake major experiments with their elementary courses because of shortage of staff and financial resources. The projects under way or completed, are at Vanderbilt University (Rendigs Fels), Carnegie-Mellon University (Phillip Saunders), and the University of Colorado (Kenneth Boulding). Reports and special materials deriving from the Vanderbilt and Carnegie-Mellon projects will be published in the *Journal of Economic Education* in the near future, and may be made more widely available to members of the profession through the Joint Council on Economic Education if this seems appropriate.

4. The Visiting Scientists Program, financed by a grant by the National Science Foundation, continued to operate effectively in 1970. A report by Professor Phillip Saunders, who is the administrator of the program under the general policy guidance of this Committee, is submitted separately.

5. The "Test of Understanding in College Economics," developed by the Joint

Council and The Psychological Corporation at the suggestion of this Committee, continues to be widely used as a research tool in studies of the effectiveness of different approaches to the teaching of elementary economics. This development is beginning to make possible for the first time significant comparisons among research projects done at different institutions. Information on the TUCE is available from the Joint Council on Economic Education.

6. The Committee has the following under active consideration:

- a. The possibility of a summer institute for graduate students, to be offered at the University of Wisconsin in 1972, devoted to developing interest in research on the effec-

tiveness of different approaches to teaching and a concern for improved undergraduate teaching.

- b. Steps to improve the labor market for economists in the junior colleges, and to provide more assistance to the junior colleges in improving the teaching of economics.

- c. Steps to help attain a more effective utilization of United States Office of Education resources in relation to research and improvement in teaching in the field of economics.

This Committee will welcome suggestions from members of the Association as to activities it should undertake, itself, or should encourage, in relation to improving the teaching of economics.

G. L. BACH, *Chairman*

REPORT ON THE VISITING SCIENTISTS PROGRAM IN ECONOMICS

Academic year 1969-70 marked the fourth year of the American Economic

The following summary figures cover the first four years of the program:

School Year	No. of Schools Visited	No. of Visit Days	Expenditure	Cash Contributions	Cost to NSF
1966-67	20	34	\$4,903	\$ 223	\$4,680
1967-68	32	43	5,274	684	4,590
1968-69	29	49½	5,596	990	4,606
1969-70	33	59	7,337	1,085	6,252

Association's participation in the Visiting Scientists Program sponsored by the National Science Foundation. The purposes of the program are: (1) to stimulate interest in economics among undergraduates in colleges and universities through visits by leaders in the profession; (2) to provide opportunities for economics staff members and administrators at the visited institutions to discuss their teaching and research problems with the visiting economists. Emphasis is placed on visits to smaller colleges and universities whose major focus in economics is at the undergraduate level.

As in the preceding year, the 1969-70 program was administered by Professor Phillip Saunders of Carnegie-Mellon University under the direction of the Association's Committee on Economic Education, Professor G. L. Bach of Stanford University, Chairman. All financial transactions were executed and recorded through the offices of the AEA's Secretary-Treasurer, Professor Harold Williamson.

The 1969-70 visits covered 24 different states from New Hampshire to Oregon and from Minnesota to Mississippi. In arranging visits to the schools selected, the informal roster of distinguished economists established in preceding years was used extensively, and it was supplemented with direct arrangements in a few specific cases. The continuing cooperation and enthusiasm of the individual economists asked to make visits is one of the most heartening aspects of the program.

Beginning in September 1970 the operation of the Visiting Scientists Program in Economics moved with Professor Saunders to Indiana University, and Rendigs Fels of Vanderbilt University became Secretary-Treasurer of the American Economic Association.

Persons wishing further information about the program are encouraged to write: Professor Phillip Saunders, Department of Economics, Indiana University, Bloomington, Indiana 47401.

PHILLIP SAUNDERS

REPORT OF THE REPRESENTATIVE TO THE NATIONAL BUREAU OF ECONOMIC RESEARCH

Staff and Conference Programs

Staff research at the National Bureau of Economic Research continued during 1970 in the following general areas: economic growth, including studies of productivity, employment and price levels, and public finance; national income, consumption, and capital formation; urban and regional economic studies; human resources and social institutions including studies related to education and the economics of the legal system; business cycles, economic forecasting, and the relation of monetary trends to income, prices, and interest rates; financial institutions and processes; economics of health; international economic relations; econometrics and measurement methods; electronic computer services in support of economic research. Brief accounts of some eighty research projects now under way are given in the National Bureau's 50th Annual Report, *Economics—A Half Century of Research, 1920–1970*, which was presented to the Board of Directors at the Annual Meeting. Copies of the report are available from the National Bureau upon request.

The Conference on Research in Income and Wealth sponsored a small "follow-up" conference on February 27, 1970, to discuss a revised and completed paper on "Business Cycle Analysis of Econometric Model Simulations," which had first been presented at the Conference on Econometric Models of Cyclical Behavior at Harvard University in November 1969. On May 21–22, 1970 a Conference on International Comparisons of Prices and Real Incomes was held at York University. A conference has been tentatively planned

for the fall of 1971 on Measurement of Economic and Social Performance.

The Universities-National Bureau Committee for Economic Research sponsored a Conference on International Mobility and Movement of Capital in Washington, D.C. on January 30–February 1, 1970. Two conferences have been tentatively scheduled in 1971, a Conference on Education as an Industry and a Conference on Secular Inflation.

The Universities-National Bureau Conference on the Application of the Computer to Economic Research held its first conference at the National Bureau on February 26, 1970. The discussion dealt with the exchange of information from economic data banks and methods of retrieval. A second conference on October 22–23, 1970 in Washington, D.C. focused on problems on the use of microdata in economic analysis. A Conference on the Role of the Computer in Economic and Social Research in Latin America is scheduled for the fall of 1971.

To commemorate its fifty years of economic research, the National Bureau held an anniversary reception and dinner at the Waldorf-Astoria Hotel in New York on February 27, 1970. George P. Shultz, then Secretary of Labor, and Professor Wassily W. Leontief of Harvard University were the featured speakers. Former Directors of Research of the National Bureau, Arthur F. Burns, Solomon Fabricant and Geoffrey H. Moore, were honored. A series of colloquia was also arranged, at which scholars and business and governmental leaders reviewed research efforts in areas of particular Bureau concern and

contemplated future research needs in the light of policy problems. Colloquia on the following topics were held during the latter part of the year: The Business Cycle Today, September 24, New York City; Finance and Money Markets, October 22, New York City; A Roundtable on Policy Issues and Research Opportunities in Industrial Organization, November 5, Chicago; Public Expenditures and Taxation, December 2, Washington, D.C.; and Economic Growth, December 10, San Francisco. A sixth colloquium on Human Capital is scheduled to be held in Atlanta in January 1971. The final event of this series will be a conference on The Future of Economic Research to be held April 23-24, 1971, in Boston. Simon Kuznets will be the principal speaker and the participants will include those who conducted the earlier colloquia.

Research Fellows appointed in 1970, their academic associations, and areas of interest were: John C. Hause, University of Minnesota, economics of education, calendar year 1970; and for the academic year 1970-71, Christopher Sims, Harvard University, econometric studies of production functions and the construction industry; and Robert Willis, Wesleyan University, economics of population.

Publications

Fourteen National Bureau publications were issued and given publicity in 1970—three volumes of conference proceedings and eleven staff reports comprised of six books and five occasional papers. The following eight publications were in press in December:

Books

New Series on Home Mortgage Yields Since 1951, by Jack M. Guttentag and Morris Beck
Analysis of Public Output, Report of Universities-National Bureau Conference, Julius Margolis, Editor

The Responsiveness of Demand Policies to Balance of Payments: Postwar Patterns, by Michael Michaely

The Ownership Income of Management, by Wilbur G. Lewellen

Price Competitiveness in World Trade, by Irving Kravis and Robert E. Lipsey

The Postwar Quality of State and Local Debt, by George Hempel

Essays on Interest Rates, Volume II, Jack Guttentag, Editor

PAPERS

Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, by Gerhard Bry and Charlotte Baschan

Meetings of the Board of Directors in 1970

A special meeting of the Board was held on April 27 to review developments in the Bureau's program. Among the items given particular attention were problems of improving social measurements, the expanding use of computers for economic research, the development of a proposed computer facility for economic research, and the feasibility of conducting some Bureau activities through branch offices.

The annual meeting of the National Bureau was held September 24, 1970. Three new Directors were elected: Atherton Bean of the International Multifoods Corporation was elected a Director at Large, succeeding Marion B. Folsom; Kelvin J. Lancaster was elected a Director by Appointment of Columbia University, succeeding Gary S. Becker; and Charles B. Reeder was elected a Director by Appointment of the National Association of Business Economists, succeeding George Cline Smith.

Officers elected at the Annual Meeting were:

Arthur F. Burns, Honorary Chairman
Theodore O. Yntema, Chairman of the Board
Walter W. Heller, Vice Chairman
John R. Meyer, President
Thomas D. Flynn, Treasurer

Douglas H. Eldridge, Vice President-Executive Secretary
Victor R. Fuchs, Vice President-Research
F. Thomas Juster, Vice President-Research
Hal B. Lary, Vice President-Research
Robert E. Lipsey, Vice President-Research
Edward K. Smith, Vice President
Joan R. Tron, Director of Publications

*History and Organization of
National Bureau*

The National Bureau was organized in 1920 in response to a growing demand for objective determination of the facts bearing upon economic problems and for their

interpretation in an impartial manner. It is a private nonprofit organization. Its Board of Directors includes men of different economic and social viewpoints with members from industry and labor, banking and finance, professional associations of economists, accountants and statisticians, and economists from university faculties. The American Economic Association is one of the representative organizations which nominate a member of the National Bureau's Board of Directors.

WILLARD L. THORP

REPORT OF REPRESENTATIVE TO THE INTERNATIONAL ECONOMIC ASSOCIATION

Three conferences were held in 1970, one in Ma'ale Hachamisha, Israel, March 30 to April 30, on "The Essence of Growth Models;" another in Liblice Castle, Czechoslovakia, May 4 to 8, on "Planning and Market Relations"; and a third in Bled, Yugoslavia, August 27 to September 2, on "The Gap between the Rich and the Poor Nations." The following economists from the United States contributed papers or comments: Peter A. Diamond, Hugh Rose, Karl Shell, and Joseph E. Stiglitz, at Ma'ale Hachamisha; Evsey D. Domar, Robert Dorfman, Richard Portes, Jeremy Siegel, and Lynn Turgeon, at Liblice Castle; and Kenneth Boulding, Jagdish Bhagwati, Hollis Chenery, Edward Fei, Simon Kuznets, Arthur Lewis, and Gustav Ranis, at Bled.

The training workshop for young research economists, originally planned for 1970, is now scheduled for Bergen, Norway, in July 1971; James R. Mirrlees, Oxford, will be in charge. A conference on "The Role of Science and Technology in Economic Development" will be held in Alpbach, Austria, from August 27 to September 2, 1971; Bruce Williams, Australia, is chairman of the program committee.

A conference on "Economic Development in South America" is being planned for December 1971; Victor L. Urquidi, Mexico, is chairman of the program committee. A conference on "Long-Term Planning and Structural Changes in a Planned Economy," will be held in Moscow, U.S.S.R., probably in the spring of 1972; Tigran S. Khatchaturov, U.S.S.R., is chairman of the program committee.

Four or five other program committees are at work to prepare for conferences in 1972 or later. Most advanced are the preparations for a conference on "Environment and Urbanization," to be held in Copenhagen in the spring of 1972, with Jerome Rothenberg, U.S., as chairman of the program committee; and for a conference on "The Place of Agriculture in the Development of Less Developed Countries," with Nurul Islam, Pakistan, as chairman of the program committee.

Several volumes of conference proceedings are at the printers or in various stages of preparation; two or three of them will be published before the end of 1971.

FRITZ MACHLUP

REPORT OF THE ECONOMICS INSTITUTE POLICY AND ADVISORY BOARD

During the year 1970 the Economics Institute has once again continued to provide an excellent transitional summer program for foreign students who have been admitted to graduate study in Economics and Agricultural Economics in the United States. Ninety-three students from 28 countries attended the session. They held admissions to 38 different United States' universities or colleges. Including this, the 13th session of the Institute, a total of 987 students, representing 79 countries, have now participated in the program prior to undertaking graduate work at 109 different universities in the United States.

The financial condition of the Economics Institute continues to be satisfactory. No problem will arise until the termina-

tion of the current grant in early 1973.

Two new members—Professors Earl Heady of Iowa State University and Professor Albert Fishlow of the University of California at Berkeley—joined the Policy Advisory Board at its Boulder meeting on December 7th. The Board discussion concentrated on three long-range items: the need for the Institute program and for additional support from the Ford Foundation beyond 1972, the continuation and further improvement of the credentials evaluation service, and possible affiliation with the American Agricultural Economics Association. As yet none of these items requires action by the AEA Executive Committee.

HENRY ROSOVSKY, *Chairman*

REPORT OF THE CENSUS ADVISORY COMMITTEE

The committee held its annual meeting on April 10, 1970, at the Bureau of the Census. Those in attendance included 12 members of the committee and 15 representatives of the census staff. The agenda included discussion of the following topics:

- a. Current developments
- b. Census programs in the 70's
- c. The 1972 economic censuses
- d. The 1970 census—progress and problems
- e. Bureau activities in the field of social indicators
- f. Second edition of *Long Term Economic Growth*
- g. Suspension of publication of advance retail trade series: criteria for reinstatement
- h. Adequacy of the construction statistics program
- i. Measurement of mergers, acquisitions, and disposals
- j. Proposal for new surveys of population change (deleted because of time)
- k. A review of the Bureau's 5-Year program proposals

In the course of its deliberations, the committee made recommendations concerning continuation of the Census/IRS link project, expansion of tabulations of industry data by establishment size, work on social indicators, publication of a new edition of *Long Term Economic Growth*, closure of statistical gaps in the housing area, endorsement of the time series program as a high priority project, additional studies of the competitive process of the U.S., recognition by the Census Bureau of the need for researching more fully the data it collects, and the interest of the committee in having any proposed changes in *Business Conditions Digest* brought before it. These recommendations will be given careful consideration by the Census staff and a full report on action taken will be made at the next annual meeting of the committee.

BERT G. HICKMAN, *Chairman*

REPORT ON THE ECONOMIC INDEX PROJECT

Publication of Volume IX of the *Index of Economic Articles in Journals and Collective Volumes, 1967*, is scheduled for December 1970. A total of nine journals was added to the 1967 volume. The items indexed were as follows: 2,850 journal articles, notes, or comments; 3,300 articles from collective volumes; and 292 books.

Double index classifications were about 30%; double author entries were about 11%. The volume for 1967 was 500 pages, about 50 pages larger than the volume for 1966.

Work is continuing on Volume X for 1968, and it is about two-thirds completed. After this volume is completed, work will continue on articles in collective volumes for the period 1960-63 (Volume VI A). At present about 20% of the work on Volume VI A has been done. The funds authorized for the Economic Index Project by the AEA Executive Committee at its meeting of December 27, 1969 will be sufficient to carry the project to its completion.

RICHARD RUGGLES, *Chairman*

FINANCIAL INSTITUTIONS AND MARKETS
Murray E. Polakoff, New York University; and
Others

640 pages / 1970 / \$12.95

MACROECONOMICS: Selected Readings

Edited by Walter L. Johnson and David R.
Karnerschen, both of the University of Missouri

405 pages / paper / 1970 / \$6.50

**PRICE THEORY IN ACTION: A Book of
Readings, Second Edition**

Edited by Donald Stevenson Watson,
George Washington University

400 pages / paper / 1969 / \$5.50

**PRICE THEORY AND ITS USES,
Second Edition**

Donald Stevenson Watson,
George Washington University

443 pages / 1968 / \$8.75

**DOMESTIC TRANSPORTATION: Practice,
Theory, and Policy, Second Edition**

Roy J. Sampson, University of Oregon and
Martin T. Farris, Arizona State University

512 pages / March 1971

ECONOMIC ESSENTIALS: A Core Approach

Roger C. Van Tassel, Clark University

410 pages / paper / 1969 / \$5.50

An Instructor's Manual is available.

**PERSPECTIVES OF ECONOMIC
DEVELOPMENT**

Salvatore Schiavo-Campo, University of
Massachusetts; and Hans W. Singer,
University of Sussex

351 pages / paper / 1970 / \$5.95

**DECISION MAKING FOR ECONOMIC
DEVELOPMENT: Text and Cases**

Gustav F. Papanek, Daniel M. Schydrowsky,
and Joseph J. Stern, all of Harvard University

163 pages / paper / 1971 / \$4.95

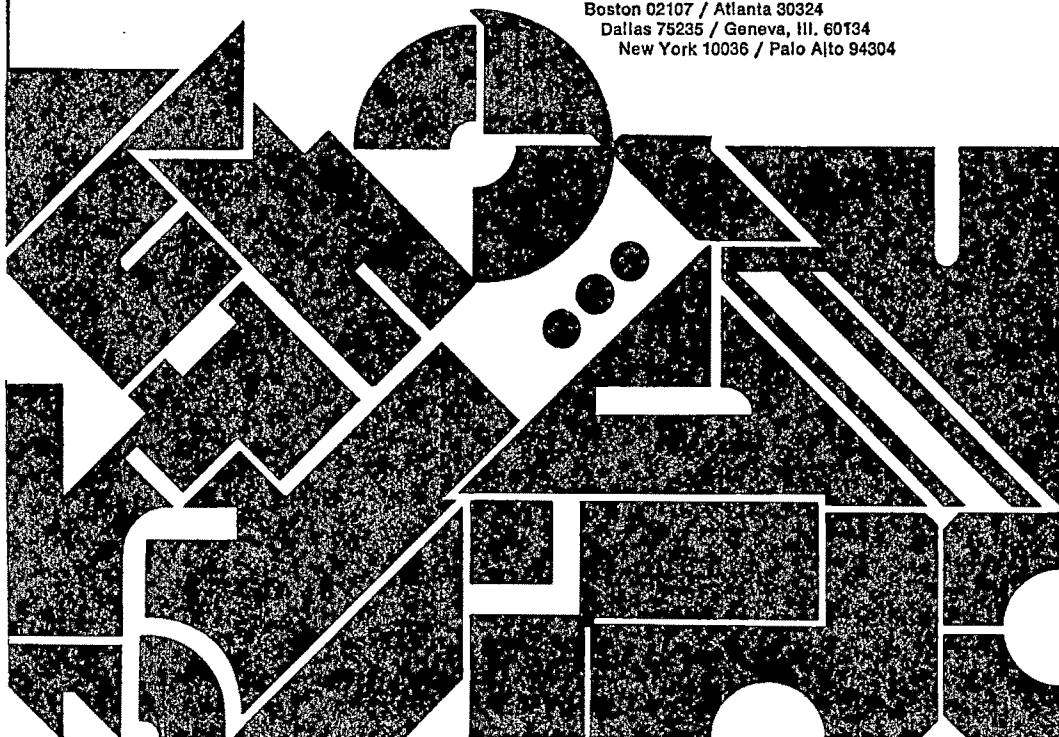
Houghton Mifflin

Publisher of The American Heritage Dictionary of the English Language

Boston 02107 / Atlanta 30324

Dallas 75235 / Geneva, Ill. 60134

New York 10036 / Palo Alto 94304



Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

The Growth Potential of the Japanese Economy

KENNETH K. KURIHARA The postwar Japanese economic "miracle" is by now familiar, but observers disagree about its causes as well as Japan's prospects for continuing growth. Professor Kurihara analyzes the fundamental forces governing the expansion of the Japanese economy to the present and for the decades ahead. His conclusions challenge some accepted assumptions about the nation's growth potential. \$8.50

The Structure of Protection in Developing Countries

BELA BALASSA and Associates Following World War II most developing countries adopted a policy of import substitution to improve their balance-of-payments positions and foster the growth of domestic industry. How successful has this approach been? This book provides criteria for evaluating the system of protection and offers guidelines for future policy formulation. \$12.00

The Invisible Resource

Use and Regulation of the Radio Spectrum

HARVEY J. LEVIN The radio spectrum is an important, but finite, economic resource. Rising demands for its use have pointed up the economic inefficiencies in the present centralized, non-price system of spectrum allocation. Professor Levin studies the economic, technical, and legal factors affecting the nation's spectrum system, examines proposals to substitute a complete market system, and suggests alternatives based on his findings. \$12.00

Society, Politics, and Economic Development

A Quantitative Approach

IRMA ADELMAN and CYNTHIA TAFT MORRIS Now in paperback. Newly revised and enlarged, this path-breaking study by two leading econometricians applies the new technique of factor analysis of socio-political aspects of economic growth in seventy-four developing countries. It has profound implications for development policy and U.S. foreign aid programs. \$8.95, \$2.95 paper

Walther Rathenau and the Weimar Republic

The Politics of Reparations

DAVID FELIX Among the factors which led from the first to the second World War, the financial reparations Germany was forced to pay may have been the most crucial. Focusing on Walther Rathenau, German spokesman for fulfillment, this political-economic study demonstrates how fulfillment was to have worked and how its failure led to World War II. \$9.00



THE JOHNS HOPKINS PRESS Baltimore, Maryland 21218

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

FRANCHISE OPERATIONS AND ANTITRUST

Donald N. Thompson, Harvard
University

An incisive study of the economic, managerial and real public policy aspects of franchising systems. Discusses the controls in franchising agreements and traces the development of franchising in several industries.

"... a very solid piece of work, a needed contribution to an understanding of the antitrust issues involved in franchising."—
Richard H. Holton, University of
California Berkeley

208 pp. June \$12.50

RESOURCES FOR URBAN SCHOOLS

Edited by Sterling M. McMurrin,
University of Utah
(Committee for Economic Development
Supplementary Paper #33)

Addresses the problem of financing and organizing urban schools for the disadvantaged. Discussion ranges from actual development aspects all the way to future prospects for this neglected issue.

ca. 140 pp. June \$10.00

CONTRACTUAL MARKETING SYSTEMS

Edited by Donald N. Thompson

A companion volume to the above, this book provides a thorough examination of the economic and legal issues of franchise marketing at both the domestic and international levels. Includes analysis of minority group capitalism, conflict resolution and the development of contractual marketing systems outside the United States.

256 pp. June \$12.50

SCIENCE, GEOPOLITICS AND FEDERAL SPENDING

Thomas P. Murphy, University of
Missouri, Kansas City

Using case studies of federal research and development agencies, this study meets head on the problem of maldistribution of Federal R & D money. Measures the importance of program administration to the geopolitical and economic development of major areas in the U.S.

ca. 256 pp. August \$15.00

INCOME PROPERTY VALUATION Principles & Techniques of Appraising Income Producing Real Estate

William N. Kinnard, Jr.,
University of Connecticut

The most comprehensive and up-to-date treatment available of income capitalization in real estate appraisal. Covers systematically the entire process of income property valuation and enables the reader to test his progress periodically with realistic review exercises at the end of each chapter.

ca. 356 pp. July \$12.95

CHOICE OF TRANSPORTATION TECHNOLOGY IN DEVELOPING COUNTRIES

Charles River Associates, Inc.

Develops and tests a model for determining the optimum type of transportation network needed by developing nations. Also studies the role of transportation facilities in the economic growth of these countries.

288 pp. August \$13.50

Heath Lexington Books

AVAILABLE

INTERNAL LABOR MARKETS AND MANPOWER ANALYSIS

Peter B. Doeringer, Harvard University
Michael J. Piore, Massachusetts
Institute of Technology

224 pp. paper \$3.95 cloth \$7.50

COBALT: AN INDUSTRY STUDY

James C. Burrows, Charles River
Associates, Inc.

256 pp. \$25.00

INDUSTRIAL LOCATION IN THE UNITED STATES

James C. Burrows, Charles T. Metcalf,
John B. Kaler, Charles River
Associates, Inc.

200 pp. \$15.00

COMPANY WAGE STRUCTURE: ISSUES AT ARBITRATION

Clifford M. Baumbach, University
of Iowa

160 pp. \$12.50

RETRAINING COURSES: ECONOMIC COSTS AND BENEFITS

Einar Hardin, Michael Borus,
Michigan State University

256 pp. \$15.00

TAX INCENTIVES

Tax Institute of America

280 pp. \$12.50

THE USE OF ECONOMETRIC MODELS BY FEDERAL REGULATORY AGENCIES

Joe L. Steele, Texas Christian
University

120 pp. \$12.50

PREDICTABILITY OF STOCK MARKET PRICES

Clive W. J. Granger, University of
Nottingham

Oskar Morgenstern, New York
University
336 pp. \$15.00

SPECULATION, HEDGING AND COMMODITY PRICE FORECASTS

Walter C. Labys, United Nations
Conference on Trade and
Development; Clive W. J. Granger,
University of Nottingham

352 pp. \$15.00

THE DEMAND FOR TRAVEL: THEORY AND MEASUREMENT

Edited by Richard E. Quandt,
Princeton University

320 pp. \$15.00

INTERNATIONAL TRAVEL— INTERNATIONAL TRADE

H. Peter Gray, Rutgers University

286 pp. \$12.50

AN ANALYSIS OF THE UNITED STATES OIL IMPORT QUOTA

James C. Burrows, Thomas A.
Domenich, Charles River Associates,
Inc.

280 pp. \$19.50

SIMULATING REGIONAL ECONOMIC DEVELOPMENT

William H. Miernyk et al., University
of West Virginia

365 pp. \$15.00

A SHORT RUN FORECASTING MODEL OF THE UNITED STATES ECONOMY

Ray C. Fair, Princeton University

274 pp. \$12.50

Heath LEXINGTON Books

D.C. Heath and Company
125 Spring Street
Lexington, Massachusetts
02173

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

new this spring from D.C. HEATH

Contemporary Economics: Problems and Policies

Willard W. Howard, Phoenix College

Edwin L. Dale, Jr., The New York Times

A teaching economist and an economic journalist combined talents to prepare this lively, well-written principles text. Using basic theoretical concepts, the authors analyze current economic problems and policies. *Instructor's Manual* available. Ready now Hardbound 544 pages

Problems in Political Economy: An Urban Perspective

David M. Gordon, National Bureau of Economic Research

Essays by leading economists and other social scientists provide a framework for the editor's basic analysis: that urban problems have developed inevitably from the structure of our economic institutions, and that these institutions must be restructured to solve the problems. Ready now Paperbound 544 pages

The American Economy in Conflict: A Book of Readings

Robert B. Carson, State University of New York at Oneonta

Both sides of timely and controversial economic issues are presented at a level suitable for beginning students. Ready now Paperbound 464 pages

Perspectives on Poverty and Income Distribution

James G. Scoville, University of Illinois

This collection of readings deals with the economic aspects of poverty in the United States. Ready now Paperbound 400 pages

Internal Labor Markets and Manpower Analysis

Peter B. Doeringer, Harvard University

Michael J. Piore, Massachusetts Institute of Technology

An innovative theory of the internal labor market is carefully developed and then used to analyze factors that influence labor economics. Ready now Hardbound/Paperbound 288 pages

Introduction to Mathematical Economics:

Matrix Algebra and Linear Economic Models

Richard H. Puckett, Board of Governors, Federal Reserve

The techniques of linear algebra are introduced and applied to economic and business problems. Ready now Hardbound 256 pages

For examination copies write:

Faculty Correspondence Manager Dept. E A



D.C. Heath and Company

125 Spring Street

Lexington, Massachusetts 02173

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Outstanding **RONALD** *books*

AGGREGATE ECONOMIC ANALYSIS

An Intermediate Exposition

Paul J. Wells, University of Illinois

Ready this Spring! This book for intermediate courses develops a concise explication of the central core of national income analysis. The aggregate economic behavior of households and businesses is carefully specified, and a model utilizing the fundamental concepts of aggregate supply and aggregate demand is developed to explain the overall performance of the economy.

Notable features are the integration of national income accounting concepts with the main body of theory, the equal importance given aggregate supply and aggregate demand in the construction of the theory, the discussion of the role of money in the behavior of the economy, and a clear exposition of the theory of investment. 1971. 289 pp. \$8.50

WAGE AND EMPLOYMENT THEORY

Sidney R. Finkel, Indiana University

Vincent J. Tarascio, University of North Carolina at Chapel Hill

This new textbook presents an economic analysis of technical, market, and institutional conditions influencing wages, employment, union organization and strategies. It begins with a critical appraisal of marginal productivity theory. On this foundation the book develops a conceptual framework of a theory of wages and introduces con-

cepts of limitational and limitative factors of production. Labor is examined as a factor of production, and union behavior is analyzed to lay the basis for discussing economic operations of trade unions and the forces influencing wage structure. Finally, there is a macroeconomic analysis of wages and employment. 1971. 242 pp. \$8.50

STATE AND LOCAL FINANCE

Edited by William E. Mitchell, University of Missouri—St. Louis
Ingo Walter, New York University

Reflecting the rapid growth in the public sector of the American economy at the state and local level, this book of readings is designed to provide a wide range of important factual and interpretive information and to add breadth and depth to teaching in the field. The twenty-nine selections, which focus on

current issues and problems, are organized in terms of sources and uses of funds. Each chapter begins with a brief introduction by the editors which relates the selections to one another and places them in the full context of the topic being discussed. 1971. 370 pp. \$8.50

The Ronald Press Company

Publishers since 1900

79 Madison Avenue • New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

for college courses . . .

SECURITY ANALYSIS AND PORTFOLIO MANAGEMENT

Henry A. Latané, University of North Carolina at Chapel Hill
Donald L. Tuttle, Indiana University

Covering the entire process of investing in stocks and bonds, book leads in a definite, systematic progression from analysis of individual securities to the ultimate objective of building the portfolio most likely to maximize long-run wealth. Whenever possible, modern decision and probability theory, statistical techniques, financial mathematics, and

the computer have been used to quantify the different phases of the investment process, test recommended procedures, and verify results. Book probes beneath surface aspects to provide understanding of the deeper running forces affecting prices, earnings, and dividends. 1970. 753 pp. \$13.50

INTRODUCTION TO MONETARY ECONOMICS

Money, Banking, and Economic Activity

Robert E. Weintraub, University of California, Santa Barbara

A rigorous and analytical discussion of money supply and money demand, monetary institutions' policy and practice, the balance of payments, and the effects of monetary development on economic activity. Book stresses theoretical issues throughout, particularly money supply theory and the Keynesian and Quantity

theories relating changes in money to other economic aggregates; provides a clear and thorough discussion of the $I, S-L, M$ model. Elementary mathematical techniques are worked carefully into the discussion as needed. 1969. 444 pp. \$9

INTERNATIONAL ECONOMICS

Theory and Policy

Ingo Walter, New York University

Emphasizing economic theory and analysis, this textbook presents the principles of international trade, payments, and commercial policy. Provides an elementary conceptual foundation of international economic analysis for students with a general economics background

in the theory of price and income determination. Throughout, extensive use is made of graphical analysis and geometric presentation of ideas. Unusually full annotated bibliographies are provided at the end of each chapter. 1968. 623 pp. \$9

The Ronald Press Company

Publishers since 1900

79 Madison Avenue • New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE

By F. M. Scherer, University of Michigan

"It's clear from the chapters I've read that INDUSTRIAL MARKET STRUCTURE is not only vastly superior to its competition but also first rate."

Professor Herbert Mohring
University of Minnesota

"I used your text in my course this summer. I and the students both found it excellent."

Professor Ben Branch
Dartmouth University



Real-world industrial markets and how they function. Isn't this what modern organization is all about? Our new text, **INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE**, by F. M. Scherer, University of Michigan, is the first really comprehensive treatment in many years—with a conceptual, verbal approach that integrates structural and behavioral models. Our text is equally at home in the undergraduate or graduate curriculum, fully accessible to any student with a basic grasp of price theory. It's designed for industrial organizations and public policy; suitable for managerial economics and price theory. It's also a "state of the art" survey for policymakers and businessmen; a fresh review of theory and application for professional economists. 1970. 550 pages. \$13.00 H/B.

INTRODUCTION TO MODERN MICROECONOMICS

By Kelvin Lancaster, Columbia University

"At present, the Lancaster work is the most innovative treatment of micro theory available for an undergraduate course. Its "linear" approach permits examination of production and consumption processes or activities, whereas the "smooth" versions of economic behavior presented in traditional price theory textbooks render such examination impossible."

Assistant Professor Harold J. Brumm, Jr., California State College, Los Angeles

The first undergraduate level microeconomics text to present an elementary account of the content of microeconomics as it is currently taught at the advanced level. All standard micro topics are covered, including supply and demand and the market, the theory of production, the theory of the firm, and consumer behavior. Emphasis is given to general equilibrium approaches, welfare economics, the problems arising from time and uncertainty, and the role of prices as information parameters. 326 pages. \$9.50 H/B.

Student Guide for Introduction to Modern Microeconomics
By David L. Schulze, Iowa State University
200 pages \$4.25 P/B.

For information about examination copies write:

College Department
RAND McNALLY & COMPANY
Box 7600 Chicago, Illinois 60680

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

New Directions in Management and Economics

A rapidly growing series bridging the gap between new developments in analytical thought and the dissemination of these ideas to a wide audience. In-depth studies of key problems reflect latest research trends set by industry, government, and leading universities. Brochure available on request.

ADVISORY BOARD: Robert Z. Aliber, *University of Chicago*; Paul R. Lawrence, *Harvard University*; Paul W. MacAvoy, *MIT*; Donald G. Marquis, *MIT*; William F. Sharpe, *Stanford University*

INVESTMENTS

New Analytic Techniques

J. Peter Williamson, *Dartmouth College*.

This book presents computer and quantitative techniques for measuring investment performance and risk, and for portfolio selection, security analysis, stock selection, and improving bond portfolios.

1971 352 pp. \$12.50

IN DEFENSE OF INDUSTRIAL CONCENTRATION

John S. McGee, *University of Washington*

1971 176 pp. \$10.00

CASH MANAGEMENT AND THE DEMAND FOR MONEY

Daniel Orr, *University of California*

1971 224 pp. \$11.75

ALASKAN OIL: COSTS AND SUPPLY

Edited by M. A. Adelman, *MIT*

1971 128 pp. \$10.75

INCOME DISTRIBUTION Facts, Theories, Policies

Jan Pen, *University of Groningen*

June, 1971 416 pp. \$12.00

ECONOMIC FORECASTING

Herman O. Stekler, *SUNY/Stony Brook*

1970 194 pp. \$11.50

Two New Paperbacks From Praeger

TECHNOLOGY AND GROWTH: The Price We Pay

E. J. Mishan, *London School of Economics and Political Science*. In this provocative book the author focuses on the current obsession with economic growth, the inept social policies that have resulted from it, and the high costs society must pay for its unchecked technological development. He shows that while consumer gadgetry is increasing, the choice of environment in which to enjoy the technology is rapidly vanishing. Consumer wants, he asserts, should be directed by beneficent administrators rather than advertising agencies and television networks, and he spells out the rationale for a new kind of policy-making. Basing his arguments on factual analysis and strong conviction, Dr. Mishan demolishes the myth of technological "progress."

1970 216 pp. \$2.95 paper; \$7.95 cloth

THE ECONOMICS OF AFRICAN DEVELOPMENT Revised Edition

Andrew M. Kamarck, *Johns Hopkins University*. Foreword by Pierre Moussa. Mr. Kamarck's widely acclaimed study of the economic development, structure, and prospects of the African countries, originally published in 1967, has been extensively revised to take account of recent events and policies, to incorporate the findings of recent scholarship on the backgrounds of national economies, and to reevaluate the potential of African nations for self-sustained economic growth. Adding his information on the experience of the last two years to his knowledge of the more distant African past, Mr. Kamarck projects the probable trends of the 1970's. The text has been expanded, and new bibliographies and maps have been provided.

1971 308 pp. \$3.95 paper; \$10.00 cloth

The College Department • PRAEGER PUBLISHERS

111 Fourth Avenue, New York 10003

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Announcing

The Bobbs-Merrill Reprint Series in ECONOMICS

*More than 325 articles from leading journals in the
field, to be available for Fall 1971 classes*

Editorial Note

The articles in The Economics Reprint Series have been selected for reprinting primarily because of their scientific interest. Many of these articles will accordingly find their audience mainly among professional-economists and more advanced students. At the same time, the Editorial Selection Committee has given articles that will be useful to students in more elementary courses preference over less readable articles of similar substantial merit. Thus, if the Series should succeed in serving all of the needs it was designed to meet, it will furnish some articles that the professional economist will want for his own reading, others that he will want to put on reading lists for his advanced students, and still others that he will want to assign in elementary courses.

The Editorial Selection Committee appreciates the help of the scores of distinguished economists in different specialties who have responded to requests to nominate articles for the Series. It regrets that the limitations on the length of the list meant that many important articles had to be left out. In a few cases copyright restrictions kept significant articles off the list. The Committee inevitably had to make many arbitrary choices and in some cases, no doubt, failed to make the best selection. Hopefully, the list can be revised and made more useful in future years. Suggestions for articles that should be added to the list are most welcome.

The articles in the Bobbs-Merrill Reprint Series in Economics are listed by author in alphabetical order. The articles are also listed separately, later in the catalog, by the special field of economics to which they belong, so that economists can more readily find the articles related to their own interests and courses. In the listings for the separate fields, there are also a large number of articles from Bobbs-Merrill Reprint Series in related disciplines. Though these latter articles were chosen for reprinting by other Editorial Selection Committees with other disciplines in mind, many of them will hopefully also prove relevant to economists. The citations of articles in the Economics Reprint Series are preceded by the designation "Econ.," and the reprints chosen by Editorial Selection Committees for other fields are similarly preceded by an abbreviated designation indicating the Reprint Series for which they were initially reprinted.

THE EDITORIAL SELECTION COMMITTEE IN ECONOMICS:

RICHARD N. COOPER, *Yale University*
ROBERT DORFMAN, *Harvard University*
JOHN F. KAIN, *Harvard University*
MANCUR OLSON, *University of Maryland*
OLIVER E. WILLIAMSON, *Wharton School
of Finance and Commerce,
University of Pennsylvania*

Complete catalogs available upon request from

The Bobbs-Merrill Company, Inc.

College Division / Indianapolis, Indiana 46268

Any reseller is free to charge whatever price it wishes for our books.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Important books, published and forthcoming

from Bobbs-Merrill:

The Politics of Taxing and Spending

by Ira Sharkansky

230 PAGES

\$3.75 PAPER

\$8.00 CLOTH

Economic Organizations and Social Systems

by Robert A. Solo

520 PAGES \$8.75 CLOTH

from Pegasus:

The Economics of the Ghetto

by Carolyn Shaw Bell

256 PAGES

\$2.25 PAPER

\$6.95 CLOTH

The Politics of Poverty

by John C. Donovan

160 PAGES

\$1.95 PAPER

\$5.75 CLOTH

Modern Britain

A Social and Economic History Since 1760

by Pauline Gregg

640 PAGES \$7.50 CLOTH

from Odyssey Press:

Labor Law and Public Policy

by Ronald A. Wykstra and Eleanor V. Stevens

480 PAGES

\$8.25 CLOTH

American Labor and Manpower Policy

by Ronald A. Wykstra and Eleanor V. Stevens

240 PAGES

\$3.75 PAPER

Information and catalogs for Bobbs-Merrill, Pegasus, and Odyssey Press are available upon request from:

The Bobbs-Merrill Company, Inc.

College Division / Indianapolis, Indiana 46268

Any reseller is free to charge whatever price it wishes for our books.

ALDINE • ATHERTON

CAN MOVE YOUR STUDENTS TO A HIGHER TEXTBOOK INDIFFERENCE CURVE!

The following titles are available on a 60-day examination basis, with charges to be dropped if the book is adopted for class use:

THE ECONOMICS OF SOCIALISM

J. Wilczynski, Royal Military College of Australia

The most comprehensive and modern source available on socialist economic systems, this book employs the latest economic data available from eight East European Socialist countries and Russia to provide students with a thorough, accurate picture of the scope of socialist economies. The author carefully analyzes the major focal points of socialist economics: planning and market, profit, production and growth, accumulation, consumption, labor, land, pricing, money and banking, fiscal policy and control, domestic and foreign trade, and international economics. 248 pp. \$8.95

INCOME DISTRIBUTION THEORY

Martin Bronfenbrenner, Carnegie-Mellon University and the University of Pittsburgh

This book is a well-grounded restatement, defense, and development of the theory of income distribution in both its micro- and macroeconomic aspects. The author, an authority in the field who spent 24 years developing the ideas in this book, balances neoclassical theories with Keynesian and Marxist approaches. He considers income distribution theory in terms of ideology, statistics, micro- and macro-economics, income policies, and the poverty problem. Approx. 576 pp. \$14.75.

EXPECTATION, ENTERPRISE AND PROFIT: The Theory of the Firm

G. L. S. Shackle, University of Liverpool

Designed for a first course in the theory of the firm, this book offers a concise treatment and excellent analysis of the major concepts in this area. Focusing on the essential institutional means of confronting the problem of uncertainty in Western economies, it concerns itself with the nature and mode of the life of a firm as it relates to policy formation. 168 pp. \$7.95

INTERNATIONAL ECONOMICS

Sidney J. Wells, Professor of Economics, University of Salford

This groundwork text details topics such as the balance of payments, workings of the IMF, the pros and cons of a fluctuating exchange rate, European unification, and the trade problems of underdeveloped countries. Although a basic undergraduate text, the discussions of recent developments in trade theory will also appeal to the expert. 330 pp. \$7.95.

ECONOMIC DIMENSIONS IN EDUCATION

Martin O'Donoghue, Trinity College, Dublin

A systematic and broad overview of the economic issues of education soundly based in original economic theory, this book is valuable reading for both students and professionals. The author examines vital topics such as the reasons for public sector activity in education, various aspects of demand and supply of education, the relationship between education and economic growth, and issues raised by different financing methods. Approx. 200 pp. \$5.95.

THE OPTIMUM QUANTITY OF MONEY AND OTHER ESSAYS

Milton Friedman, University of Chicago

These essays by one of the world's foremost monetary economists apply a single coherent view of the role of money to a wide variety of specific topics in monetary theory, the empirical analysis of monetary phenomena, and the conduct of monetary policy. This collection—essential reading for all professionals and graduate students in the field—presents the most comprehensive view available of the thinking developed at the leading center of monetary analysis. 296 pp. \$9.75.

PUBLIC EXPENDITURE

Jesse Burkhead and Jerry Miner, Syracuse University

In all highly industrialized countries, public expenditure equals or exceeds private sector expenditures. This is the first book to survey and interpret a field of new research in public expenditure as a function of urbanism. The authors integrate positive theory and empirical analysis about the subject, concentrating on the questions regarding public goods rather than the traditional economics of taxation. Approx. 352 pp. \$12.95.

PUBLIC FINANCE

Carl Shoup, Columbia University

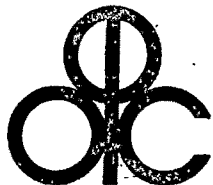
"A very scholarly book of genuine value to its field by Shoup, one of the outstanding authorities in public finance in the world at the present time."—Choice
A major new treatise on the theory and practice of public finance, this text is unique in its presentation of a worldwide perspective, in its treatment of both the instruments and goals of public finance, and in its detailed analysis of new issues. 660 pp. \$12.95.

CAUSAL MODELS IN THE SOCIAL SCIENCES

edited by Hubert M. Blalock, Professor of Sociology, University of North Carolina

Causal models—formal theories stating the relationships between precisely specified variables—have become an increasingly important tool of the social scientist. This collection of 27 articles, 8 of them previously unpublished, stands alone as a source book on the causal modeling approach to theory construction and data analysis. Approx. 512 pp. Clothbound, \$15.00. Paperbound, \$9.75.

ALDINE ATHERTON 529 South Wabash Avenue, Chicago, Ill. 60605



overseas development council offers
A NEW MONOGRAPH SERIES

DEBT AND TERMS OF AID

Charles R. Frank, Jr.

A study of the growing foreign debt burden of the developing nations: its dimensions, its causes, and how it might be relieved. (March 1970, 51 pp., \$1.00)

JOBS AND AGRICULTURAL DEVELOPMENT

Robert d'A. Shaw

A study of the effects of a new agricultural technology on employment in poor nations. (December 1970, 75 pp., \$1.00)

AMOUNT AND SHARING OF AID

Jagdish N. Bhagwati

A study of the levels, terms, conditions, and distribution of development assistance. (August 1970, 197 pp., \$1.00)

TRADE FOR DEVELOPMENT

Harald B. Malmgren

An examination of the trade policies of rich and poor countries and their effects on economic development. (January 1971, 90 pp., \$1.00)

Orders for single copies of individual titles will be mailed postpaid if check accompanies order. The discount orders of a single title in the Monograph series is: 10% for 11-50 copies; 15% for 51-100 copies; and 20% for more than 100 copies; plus postage or shipping costs.

Overseas Development Council • 1717 Massachusetts Avenue, N.W. #501 • Washington, D. C. 20036

NEW

BUSINESS AND SALES TAXATION IN THE REPUBLIC OF IRELAND

**A Study of Tax Incentives as a Tool
for Economic Development**

by

JOSEPHINE CORRIGAN, PH.D.

Economics Department

St. John's University

This book comments on the problems of development economics, with special emphasis on taxation for economic growth. The author devotes the greater part of the book to the taxes most important to business enterprises. It is the only large-scale study of Irish national taxation written on either side of the Atlantic. This book should prove useful to many American readers, including students of the Irish economy, of economic development, and American businessmen who are established in Ireland or contemplating expansion there.

200 pages

\$8.00

Also available:

Handbook of Selected Sample Surveys in the Federal Government: With Annotated Bibliography, by Jean Namias, Professor of Economics, St. John's University. **\$12.50**

The Management of Forecasting: Readings. Contributors include Julius Shiskin, U.S. Bureau of the Census; Charles Yang MSU; and Jean Namias, St. John's University. **\$9.50**

For further information,

write . . . **St. John's University Press**
Jamaica, New York 11432

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers